



This is a repository copy of *Evaluation of the performance of Euro-CORDEX Regional Climate Models for assessing hydrological climate change impacts in Great Britain: A comparison of different spatial resolutions and quantile mapping bias correction methods.*

White Rose Research Online URL for this paper:  
<http://eprints.whiterose.ac.uk/161436/>

Version: Accepted Version

---

**Article:**

Pastén-Zapata, E., Jones, J.M. [orcid.org/0000-0003-2892-8647](https://orcid.org/0000-0003-2892-8647), Moggridge, H. et al. (1 more author) (2020) Evaluation of the performance of Euro-CORDEX Regional Climate Models for assessing hydrological climate change impacts in Great Britain: A comparison of different spatial resolutions and quantile mapping bias correction methods. *Journal of Hydrology*, 584. 124653. ISSN 0022-1694

<https://doi.org/10.1016/j.jhydrol.2020.124653>

---

Article available under the terms of the CC-BY-NC-ND licence  
(<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

1 Title:

2 Evaluation of the performance of Euro-CORDEX Regional Climate Models for assessing hydrological  
3 climate change impacts in Great Britain: a comparison of different spatial resolutions and quantile  
4 mapping bias correction methods

5 Authors:

6 Ernesto Pastén-Zapata<sup>1,2\*</sup>, Julie M. Jones<sup>2</sup>, Helen Moggridge<sup>2</sup>, Martin Widmann<sup>3</sup>

7 Affiliations:

8 <sup>1</sup> Geological Survey of Denmark and Greenland (GEUS), Department of Hydrology, Denmark

9 <sup>2</sup> Department of Geography, University of Sheffield, United Kingdom

10 <sup>3</sup> School of Geography, Earth and Environmental Sciences, University of Birmingham, United Kingdom

11

12 \* Corresponding author:

13 Ernesto Pastén-Zapata  
14 Department of Hydrology  
15 Geological Survey of Denmark and Greenland (GEUS)  
16 Oester Voldgade 10  
17 1350 Copenhagen K  
18 Denmark  
19 epz@geus.dk  
20

21 Abstract

22 Regional Climate Models (RCMs) are an essential tool for analysing regional climate change impacts,  
23 such as hydrological change, as they provide simulations with more small-scale details and expected  
24 smaller errors than global climate models. There has been much effort to increase the spatial resolution  
25 and simulation skill of RCMs (i.e. through bias correction), yet the extent to which this improves the  
26 projection of hydrological change is unclear. Here, we evaluate the skill of five reanalysis-driven Euro-  
27 CORDEX RCMs in simulating precipitation and temperature, and as drivers of a hydrological model to  
28 simulate river flow on four UK catchments covering different physical, climatic and hydrological  
29 characteristics. We use a comprehensive range of evaluation indices for aspects of the distribution such  
30 as means and extremes, as well as for the structure of time series. We test whether high-resolution RCMs  
31 provide added value, through analysis of two RCM resolutions, 0.44° (50 km) and 0.11° (12.5 km), which  
32 are also bias-corrected employing the parametric quantile-mapping (QM) method, using the normal  
33 distribution for temperature, and the Gamma (GQM) and Double Gamma (DGQM) distributions for  
34 precipitation. The performance of these is considered for a range of meteorological variables and for the  
35 skill in simulating hydrological impacts at the catchment scale.

36 In a small catchment with complex topography, the 0.11° RCMs clearly outperform their 0.44° version for  
37 precipitation and temperature, but when used in combination with the hydrological model, fail to capture  
38 the observed river flow distribution. In the other (larger) catchments, only one high-resolution RCM  
39 consistently outperforms its low-resolution version, implying that in general there is no added value from  
40 using the high-resolution RCMs in those catchments. Both resolutions produce river flow simulations that  
41 cover the observed flow duration curve, but the ensemble spread is large and therefore the simulations

42 are difficult to use in practice. GQM decreases most of the simulation biases, except for extreme  
43 precipitation and high flows, which are further decreased by DGQM, which also reduces the multi-model  
44 simulation spread. Bias correction does not improve the representation of daily temporal variability  
45 measured by the Nash-Sutcliffe Efficiency Index, but it does for monthly variability, in particular when  
46 applying DGQM, which reduces most of the simulation biases. Overall, an increase in RCM resolution  
47 does not imply a better simulation of hydrology and bias-correction represents an alternative to ease  
48 decision-making.

49

## 50 **1. Introduction**

51 Global General Circulation Models (GCMs) are the main tool for climate change projections. However,  
52 their spatial resolution is usually not finer than 100 km x 100 km (Rummukainen, 2016), limiting their skill  
53 to simulate local climate. Regional Climate Models (RCMs) focus on specific subcontinental or subnational  
54 domains, incorporating regional features such as topography, coasts and islands more accurately.  
55 Consequently, RCMs improve the simulation of small-scale processes that affect precipitation, such as  
56 orographic forcing (Rummukainen et al., 2015; Di Luca et al., 2015), and are expected to yield more  
57 accurate projections of climate change at finer spatial scales. RCMs have been used extensively to evaluate  
58 the impacts of climate change on hydrology, such as changes in mean river flow, floods or low flows (e.g.  
59 Kay et al., 2015; Kay and Jones, 2012; Mendoza et al., 2016; Teng et al., 2015; Prudhomme et al., 2013;  
60 Cloke et al., 2013).

61 The resolution of RCMs has increased over time with the availability of higher computer power.  
62 Currently, the spatial resolution of RCMs varies from 50 km x 50 km to less than 5 km x 5km (Rummukainen,  
63 2016; Rockel et al., 2015). Due to their increased representation of regional features and small-scale  
64 processes, RCMs generally simulate the current regional climate better than their driving data (Feser et al.  
65 2011; Di Luca et al., 2015). Nevertheless, this might not be true in regions mainly influenced by large-scale  
66 climatic processes (Eden et al., 2014). Therefore, the added value of high-resolution RCMs depends on  
67 the analysed region, variable and context (Rummukainen, 2016).

68 An important driver for increasing RCM resolution is the need to improve the analysis of climate change  
69 impacts for decision-making (e.g. Macadam et al., 2016; Qian et al., 2015). For hydrology, the standard  
70 analysis of climate change impacts generally involves coupling uncorrected or bias-corrected GCM or RCM  
71 precipitation and temperature outputs with hydrological models to simulate river flow scenarios (e.g.

72 Teutschbein and Seibert; 2012; Huang et al., 2014; Teng et al., 2015). In Great Britain, these studies focus  
73 on one (or more) of four main topics: 1) the contribution of the GCMs, RCMs, emission scenarios and bias-  
74 correction techniques to the uncertainty of the change projection (e.g. Prudhomme and Davies, 2009; Kay  
75 et al., 2009; Arnell, 2011; Christerson et al., 2012), 2) the impact of the bias correction techniques on the  
76 projections (e.g. Prudhomme et al., 2013; Cloke et al., 2013; Wetterhall et al., 2012; Kim et al., 2016), 3)  
77 projections of future floods (Cloke et al., 2013; Kay et al., 2015; Wetterhall et al., 2012; Kay and Jones,  
78 2012), and, 4) projections of future low flows (Wilby and Harris, 2006; Arnell, 2011; Fowler and Kilsby,  
79 2007).

80 Some studies have identified a consistent improvement in hydrological simulation skill with increasing  
81 RCM resolution for the annual mean river flow (Huang et al. 2014). For the simulation of river flow peaks  
82 as a response to extreme precipitation events, previous studies found no improvement when increasing the  
83 model resolution (Kay et al. 2015; Huang et al., 2014). Others studies found that the improvement depends  
84 on the catchment size and on the evaluation index (Dankers et al. 2007), whilst others found an  
85 improvement when simulating seasonal flow and hydrologic signatures aimed to represent diverse  
86 hydrologic processes (e.g. runoff ratio, center time of runoff) (Mendoza et al., 2016). However, these studies  
87 have only used one RCM to perform the comparison as, to date, there has been no systematic study using  
88 a large number of RCM simulations to test the effect of RCM resolution on hydrological simulation skill.

89 The first aim of this paper is to use the EURO-CORDEX simulations (Jacob et al. 2014) to robustly  
90 assess the added value of increasing RCM resolution on hydrological simulations. The Euro-CORDEX  
91 simulations at  $0.11^\circ$  (12.5 km x 12.5 km) and  $0.44^\circ$  (50 km x 50km) have the same lateral boundaries and  
92 the parameterisations of each RCM are the same at both resolutions, thus making them ideal for such a  
93 comparison. This work builds on assessments of the  $0.11^\circ$  and  $0.44^\circ$  Euro-CORDEX RCMs at reproducing  
94 observed temperature and precipitation distributions, including extremes and dry/wet spell lengths. Results  
95 vary among the studies. Some found a higher accuracy for the  $0.11^\circ$  RCMs for Europe when evaluating the  
96 mean and extreme precipitation at a daily and sub-daily temporal resolution (Prein et al. 2015, Fantini et al  
97 2016), whereas others did not find an improvement in accuracy when assessing the spatio-temporal  
98 patterns of the monthly and seasonal precipitation and temperature (Kotlarski et al. 2014). For the Alps

99 Torma et al. (2015) found a higher skill for the 0.11 RCMs when simulating the spatial distribution of the  
100 mean, extreme and intensity of precipitation, while Casanueva et al. (2016) showed for the Alps and Spain  
101 that the best performance depends on the RCM, season and validation index when evaluating precipitation  
102 intensity, frequency, mean and extremes.

103 Biases in RCM simulations are due to parameterisation of sub-grid processes, limited representation  
104 of local features, incorrect boundary conditions and differences between spatial resolutions of the  
105 simulations and observations (Ehret et al., 2012; Benestad, 2010). Therefore, RCMs require post-  
106 processing for many applications (Christensen et al., 2008). Statistical bias-correction techniques reduce  
107 biases in the mean, variance or the complete distribution of simulated climate variables (reviews in Maraun  
108 et al., 2010; Teutschbein and Seibert, 2012; Maraun and Widmann, 2018; Lafon et al., 2013). Quantile  
109 mapping (QM) is one of the standard techniques used (Piani et al., 2010; Teutschbein and Seibert, 2012;  
110 Maurer et al., 2014). Whilst effective, bias correction has important limitations that are further discussed in  
111 the conclusions.

112 To date, a detailed comparison of the simulation skill of bias-corrected high- and low-resolution model  
113 outputs for aspects that are important for hydrological studies (e.g. means, extremes, daily sequence) has  
114 not been undertaken. The second aim of this study addresses this research gap by conducting a detailed  
115 evaluation of aspects that are relevant for the hydrological regime such as seasonal precipitation,  
116 occurrence of extreme events, and monthly and daily pairwise indices (assess the skill to reproduce the  
117 observed time-series). The evaluation of these aspects allows identifying the capabilities and weaknesses  
118 of the impact assessments. Here, the simulations are evaluated against gauged data, working as a mean  
119 to assess the plausibility of the simulation outputs using uncorrected and bias-corrected RCMs. This work  
120 builds on studies that have assessed climate variables. For instance, the bias-corrected Euro-CORDEX  
121 simulations, at both resolutions, have a similar skill at capturing the wet-day intensity and precipitation  
122 frequency (Casanueva et al., 2016).

123 Here, we therefore address the two above-mentioned research aims by evaluating the simulation skill  
124 of five uncorrected and bias-corrected Euro-CORDEX RCMs at 0.11° and 0.44° using a range of  
125 temperature, precipitation and river flow indices, evaluating the mean along with high and low extremes,

126 frequency of occurrence and daily and monthly simulation sequence. By using a multi-model ensemble,  
127 this analysis provides a robust understanding of the added value of high-resolution simulations and post-  
128 processing approaches for hydrological impact studies. We analyse four diverse catchments across Great  
129 Britain, representative of different climate and physical characteristics, focusing on the following questions:

- 130 1) Based on a range of selected indices, is the performance of the 0.11° Euro-CORDEX RCMs better  
131 than their 0.44° version to simulate (a) climate and (b) river flow?
- 132 2) Is the current skill of the Euro-CORDEX RCMs sufficient to generate plausible inputs for the  
133 analysis of climate change impacts on hydrology and how does this compare to the inputs from  
134 bias-corrected simulations?
- 135 3) Is there any improvement in the simulation skill of precipitation and river flow when using a Double  
136 Gamma Quantile Mapping (DGQM) bias correction compared to the usual Gamma Quantile  
137 Mapping (GQM) approach?

138 Given the associated computational cost (Bucchignani et al., 2016) and the potential for improving the  
139 skill of climate simulations, especially for impact assessments (Ehret et al., 2012), there is a clear need for  
140 rigorous evaluation of the added value of increasing RCM resolution. Previous hydrological impact studies  
141 have analysed this issue using one or two RCMs (e.g. Mendoza et al., 2016; Kay et al., 2015). However,  
142 their results might not be transferable to other RCMs, as each has its own parameterisations.

143 GQM inflates the precipitation extremes, producing unreliable flood simulations. Whilst this is a known  
144 issue (Cloke et al., 2013; Huang et al., 2014), no study has exhaustively compared the results between  
145 using the GQM and the DGQM approaches using extreme indices. This study provides a comprehensive  
146 analysis of such gaps.

## 147 **2. Data and method**

### 148 **2.1. Observation databases and study catchments**

149 The observations are used to calibrate the hydrological model (Section 2.2), develop the bias  
150 correction method (Section 2.3) and to compare the outputs of the RCMs to evaluate their simulation skill  
151 (Section 2.5). We employ gridded observations based on weather stations, as these are better comparable  
152 to the outputs of the climate models which produce an areal average for each gridbox (following Osborn

153 and Hulme, 1997). We use the Centre for Ecology and Hydrology (CEH) Gridded Estimates of Areal Rainfall  
154 (CEH-GEAR) dataset (Tanguy et al. 2014) as 1km x 1km gridded daily precipitation observations (Keller et  
155 al., 2015). Records from the Natural Environment Research Council (NERC) Hydrology and Ecology  
156 Research Support System (CHESS) (Robinson et al., 2017a, 2017b) are used as 1 km x 1 km gridded daily  
157 temperature observations. The 1 km x 1 km gridded CHESS-PET dataset is employed as potential  
158 evapotranspiration (PET) observational reference. CHESS-PET uses the Penman-Monteith equation  
159 (Monteith, 1965) to calculate daily PET using climate variables from the Met Office Rainfall and Evaporation  
160 Calculation System (MORECS) (Hough and Jones, 1997) as input. All these datasets cover the period 1961  
161 to 2010. A detailed description of the methodology and weather stations used to develop the gridded  
162 datasets can be found in Robinson et al. (2017a, 2017b) and Tanguy et al. (2014). We use river flow  
163 observations from the CEH's National River Flow Archive (NRFA). The available river flow observations for  
164 the 1961-2010 period varies in each catchment, with a minimum of 30 years of continuous records.

165 We analyse four catchments within the UK. The catchments have long river flow records and cover  
166 regions that are representative of the different climate and catchment types that can be found within the  
167 UK. These are the Upper Thames, Glaslyn, Calder and Coquet catchments (Fig. 1). This set of catchments  
168 with different characteristics (Table 1) can aid identifying key features that impact on the simulation skill of  
169 the RCMs. The smallest catchment is the Glaslyn, which has the most complex topography and highest  
170 rainfall. The largest catchment is the Upper Thames (1616 km<sup>2</sup>), which also has the least complex  
171 topography. The Calder and Coquet are intermediate in terms of area, elevation and precipitation. These  
172 catchments have been studied before using bias-corrected climate projections (QM, normal distribution for  
173 temperature and Gamma distribution for precipitation) from the HadRM3-PPE RCM (Prudhomme et al.,  
174 2013).

## 175 **2.2. RCMs**

176 We evaluate two spatial resolutions (0.11° equivalent to 12.5 km x 12.5 km and 0.44° equivalent to  
177 50 km x 50 km) of five Euro-CORDEX RCMs (Jacob et al., 2014) driven by the ERA-Interim reanalysis (Dee  
178 et al., 2011), the so-called 'evaluation simulations'. The evaluation simulations are used as these are driven  
179 by observations and consequently simulate the internal variability in synchronicity with reality, in contrast

180 to the historical simulations. The assessed RCMs are shown in Table 2 (refer to Table 1 in Kotlarski et al.  
 181 (2014) and Table 1 in Prein et al. (2015) for a detailed RCM description). These models are selected as  
 182 they have the best performance to reproduce observations in the British Isles according to Kotlarski et al.  
 183 (2014). When more than one RCM cell is needed to fully cover the catchment we use the mean of the cells  
 184 to represent the catchment's climate simulations (see Fig. 1).

185 **2.3. Bias correction**

186 QM is used based on parametric representations of the simulated and observed distributions (Piani  
 187 et al., 2010). For each month of the year, the Gamma distribution is fitted to the observed and simulated  
 188 gridded daily precipitation and the normal distribution to the observed and simulated gridded daily  
 189 temperature. RCMs generally simulate too many days with very low precipitation and not enough dry days.  
 190 Therefore, in an initial step the QM method adjusts the number of simulated dry days in the RCM evaluation  
 191 simulations such that they match with the number of observed dry days by including a wet day threshold  
 192 and replacing all values below it with zero. After the wet-day adjustment, the distributions of the simulations  
 193 and observations are matched using their cumulative distribution functions (CDF). The method is  
 194 represented by the following equations:

195 
$$P_c(t) = F_g^{-1}(F_g(P_R(t), \alpha_R, \beta_R), \alpha_O, \beta_O) \quad (1)$$

196 
$$T_c(t) = F_n^{-1}(F_n(T_R(t), \mu_R, \sigma_R^2), \mu_O, \sigma_O^2) \quad (2)$$

197  
 198 Where  $P_c(t)$  and  $P_R(t)$  represent the bias-corrected and raw RCM daily precipitation, respectively.  
 199 Likewise,  $T_c(t)$  and  $T_R(t)$  stand for the bias-corrected and raw RCM daily temperature. The raw RCM CDF  
 200 is symbolized with  $F$ , and  $F^{-1}$  stands for the observations inverse CDF. The 'g' and 'n' subscripts represent  
 201 the Gamma and normal distributions, respectively. The precipitation shape and scale parameters are  
 202 symbolised by  $\alpha$  and  $\beta$  and the temperature mean and standard deviation by  $\mu$  and  $\sigma$ , respectively. Finally,  
 203 the 'R' and 'O' subscripts are used to symbolize the distribution parameters from the raw RCM and  
 204 observations, respectively.

205 GQM focuses on the most frequent values (e.g. means) (Teng et al., 2015; Yang et al., 2010).  
 206 Consequently, the corrected precipitation extremes tend to be inflated compared to the observations



207 (Cannon et al., 2015). Therefore, we also bias-correct precipitation using the DGQM. The methodology is  
208 mainly the same as the GQM with the difference that the simulated precipitation distribution is divided in  
209 two segments. Each is corrected separately, generating correction parameters for each section. In our  
210 study, the distribution is divided at the 90<sup>th</sup> percentile because at this percentile the biases inflate (see  
211 section 3.2.2.1).

212 For the 0.11° RCMs, the spatial scale of the simulations and the observations are approximately the  
213 same and the method can be viewed as a pure bias correction. In contrast, the output of the 0.44° is given  
214 on a larger scale than the observations and thus the QM also includes a downscaling aspect to account for  
215 the difference in distributions on different spatial scales. We note that due to the existence of sub-grid  
216 variability QM is in principle problematic as the corrected values for all sub-grid locations would have  
217 unrealistic high correlations (Maraun, 2013). However, this limitation is not of high relevance for our study  
218 as we bias-correct the distributions for the entire catchments.

#### 219 **2.4. Hydrological simulation**

220 The Hydrological Modeling System from the US Army Hydrologic Engineering Center (HEC-HMS)  
221 (Scharffenberg, 2013) is used to simulate the catchments' daily river flow. HEC-HMS has been successfully  
222 used before to analyse climate change impacts on water resources in other regions (e.g. Babel et al., 2014;  
223 Azmat et al., 2015). An advantage of the model is the available guidance for the estimation of parameters.  
224 Here, the model is run using its continuous, lumped arrangement. Observed precipitation and PET time  
225 series are used as input for the calibration and validation of the model. Afterwards, the raw and bias-  
226 corrected RCM simulations drive the model to generate the river flow simulations.

227 Evapotranspiration controls the moisture returning from the Earth's surface to the atmosphere and  
228 therefore impacts on the river flow. PET estimates the amount of water returning to the atmosphere when  
229 enough water is present in the surface of the catchment. Climate models do not simulate PET directly, thus  
230 it is estimated indirectly with formulas using variables from the climate models as input. There is no  
231 consensus on whether temperature-based or physically-based formulas provide better results in a climate  
232 change context (Kay et al., 2013) as the data required by the physically-based formulas is uncertain in the  
233 climate model simulations compared to the input from one variable formulas (Kingston et al., 2009). This

234 has been discussed and explored elsewhere (please refer to: Seiller and Anctil, 2016; Kingston et al., 2009;  
 235 Kay and Davies, 2008; Kay et al., 2013). We estimate PET using the Oudin formula (Oudin et al., 2005) as  
 236 it has given accurate results before (e.g. Oudin et al., 2005; Kay and Davies, 2008).

$$237 \begin{cases} PET \text{ (mm day}^{-1}\text{)} = \frac{R_e}{\lambda\rho} \left( \frac{T+5}{100} \right) & \text{if } T + 5 > 0 \\ PET \text{ (mm day}^{-1}\text{)} = 0 & \text{otherwise} \end{cases} \quad (5)$$

238 The extraterrestrial solar radiation ( $R_e$ ) is the solar radiation received at the top of the Earth's  
 239 atmosphere which can be estimated by the latitude and day of the year. The density of water is symbolized  
 240 by  $\rho$ , the latent heat flux by  $\lambda$  (2.45 MJ/kg) (Allen et al., 1998) and  $T$  is the daily mean temperature ( $^{\circ}\text{C}$ ).  
 241 When driven by observed temperature, the Oudin formula gave results similar to the CHESS-PET dataset  
 242 for 1973 to 2010 (Pasten-Zapata, 2017).

## 243 2.5. Hydrological model calibration

244 The hydrological model is calibrated and validated against the observations using a split sample  
 245 test. Considering the available uninterrupted daily river flow records, for each catchment two same-length  
 246 non-overlapping time periods are used: one for calibration and the other for validation. The period with  
 247 available river flow observations varies for each catchment. The period with observations for each  
 248 catchment is selected and divided into two equal-length, non-overlapping periods. Calibration is done for  
 249 the more recent period and validation for the other portion of the sample. Three indices are assessed: the  
 250 low flows simulation is evaluated using the Q95 (flow equalled or exceeded 95% of the time), the high flows  
 251 by the Q10 (flow equalled or exceeded 10% of the time) and the Nash-Sutcliffe Efficiency Index (NSE)  
 252 which evaluates the fit of the simulated and observed river flow. The NSE ranges from 1 (perfect fit) to  
 253 negative (unreliable model) (Nash and Sutcliffe, 1970). In the NSE formula,  $Q_t^{obs}$  and  $Q_t^{sim}$  stand for the  
 254 observed and simulated river flow at time step  $t$ , respectively.  $Q^{mean}$  is the average of the observed river  
 255 flows during the complete period.

$$256 NSE = 1 - \left[ \frac{\sum_{t=1}^n (Q_t^{obs} - Q_t^{sim})^2}{\sum_{t=1}^n (Q_t^{obs} - Q^{mean})^2} \right] \quad (6)$$

## 257 **2.6. RCM validation approach and indices**

258 Validation is important to assess the RCM simulation skill before and after bias correction (Eden et  
259 al., 2014). Here, a five-fold cross-validation approach is used: 1) the study period is divided into five same-  
260 length, non-overlapping blocks, and 2) the QM methods are trained using four blocks and the remaining  
261 block is corrected using the parameters from the training period (Maraun et al., 2015). The corrected blocks  
262 are concatenated to time series for the entire period from which the performance measures for the bias-  
263 corrected precipitation and temperature are derived.

264 A range of distribution-based and time series-based indices evaluate the skill of the raw and bias-  
265 corrected RCM outputs to simulate the observed precipitation, temperature and river flow. The indices  
266 assess biases in the means, low and high extremes, inter- and intra-annual variability and correlations for  
267 each variable (see Table 3). RCMs are then ranked based on their skill to simulate all indices relative to the  
268 skill of the other RCMs at both resolutions. As we are evaluating the outputs of 10 RCMs (5 high-resolution  
269 and 5 low-resolution), each RCM is given a value between 1 (best) and 10 (worst) based on their simulation  
270 skill. Thus, simulation skill refers to the biases present in the models compared to the available  
271 observations considering all the metrics used in this study. We use the complete time series (dry days  
272 included) to estimate the precipitation indices. Even when driven by “perfect boundary conditions”, a close  
273 similarity between the RCM simulations and observations is not expected (Kay et al., 2015) due to subgrid  
274 variability or internal variability because the boundary conditions do not fully determine the weather states  
275 within the RCM. Nevertheless, we include daily and monthly pairwise indices as these are important for the  
276 river flow simulation. We left out the hydrological model uncertainty source intentionally to solely evaluate  
277 the effects of increasing RCM resolution. Thus, we compare the river flow simulations driven by RCM  
278 outputs against the river flow simulations driven by the observed temperature and precipitation.

## 279 **3. Results**

280 This section begins by showing hydrological model simulation skill followed by the evaluation of the  
281 simulation skill of the uncorrected RCMs for temperature, precipitation and river flow. Finally, we compare  
282 the biases that remain after bias-correcting precipitation using the GQM and DGQM and their impacts on  
283 the river flow simulation.

### 284 **3.1. Calibration and validation of the hydrological model**

285 Firstly we evaluate the hydrological model simulation skill using climate observations as input.  
286 Depending on the catchment, the length of the overall evaluation period ranges from 34 to 49 years. The  
287 daily NSE varies between 0.62 (Calder) and 0.78 (Glaslyn) for calibration and between 0.52 (Coquet) to  
288 0.78 (Glaslyn) for validation (Table 4). These results indicate a moderate to good simulation skill overall  
289 compared to the NSE values from similar studies which vary from 0.45 to 0.9 (e.g. Arnell, 2011; Walsh et  
290 al., 2015; Cloke et al., 2013). The Q10 bias ranges between -6% and 11% for the calibration and between  
291 -5% and 7% for the validation. Similarly, the Q95 bias ranges between -27% and -11% for the calibration  
292 and between -44% and 6% for the validation. Overall, the simulation of high flows is very good and moderate  
293 to very good for the low flows. More detail on the calibration and validation results can be found in the work  
294 from Pasten-Zapata (2017).

### 295 **3.2. Evaluation of the RCM simulation skill**

296 We now assess the skill of the RCMs at simulating climate and river flow, firstly for the raw  
297 simulations and then for the bias-corrected outputs. We only show robust results for the analysis of the  
298 indices (e.g. if all RCMs from a particular resolution underestimate or overestimate an index). We also  
299 evaluate the multi-model percentile bias for each variable and use a skill rank to enable comparison of the  
300 different RCMs over the different performance indices. The ranking is only estimated for the uncorrected  
301 simulations as the biases after the correction are small and similar among the RCMs. Thus, ranking the  
302 bias-corrected simulations would give meaningless results.

#### 303 **3.2.1. Uncorrected RCM simulations**

##### 304 **3.2.1.1. Temperature**

305 We begin with assessing the ability of the RCMs to simulate temperature. The 0.11° RCMs  
306 underestimate the annual mean temperature for the upper Thames (Fig. 2a, ii), Calder (Fig. 2c, ii) and the  
307 Coquet (Fig. 2d, ii) catchments, whereas the 0.44° RCMs overestimate the annual mean temperature for  
308 the Glaslyn (Fig. 2b, ii) and Coquet (Fig. 2d, ii) catchments. The monthly mean temperature bias for the  
309 0.11° RCMs is larger for the Calder (between and 0.5 °C and 1.1 °C) (Fig. 2c, ii) and smaller for the Glaslyn  
310 catchment (between 0.4 °C and 0.7 °C) (Fig. 2b, ii). In contrast, the monthly mean temperature bias of the

311 0.44° RCMs is larger for the Glaslyn (between 0.4 °C and 1.2 °C) (Fig. 2b, ii) and smaller for the Calder  
312 catchment (between 0.8 °C and 1.0 °C) (Fig. 2c, ii).

313 We use the simulation spread to evaluate the simulation skill of each resolution. The spread  
314 represents the range between the highest and lowest simulated value considering all RCMs at each  
315 resolution and all gridcells within a catchment. The temperature percentile bias spread for the upper  
316 Thames is similar for both resolutions except between the 40<sup>th</sup> and 60<sup>th</sup> percentile where the 0.44°  
317 simulation include larger positive biases (Fig. 3a). For the Glaslyn catchment, the 0.44° simulations  
318 overestimate temperature for almost all percentiles, while the biases of the 0.11° simulations are smaller  
319 (Fig. 3b). For the Calder catchment, the 0.44° RCM spread includes the no bias threshold for all percentiles,  
320 whereas the 0.11° RCMs underestimate temperature between the 40<sup>th</sup> and 90<sup>th</sup> percentile (Fig. 3c). Finally,  
321 in the Coquet catchment the 0.44° simulations overestimate temperature below the 70<sup>th</sup> percentile and the  
322 0.11° simulations underestimate it between the 40<sup>th</sup> and 80<sup>th</sup> percentiles (Fig. 3d). The Pearson correlation  
323 coefficients of the daily time series vary between 0.91 and 0.97 in all catchments for both resolutions (Figs.  
324 2, iii).

325 Integrating the RCM simulation skill of all the indices into a ranking shows that, in the upper  
326 Thames, two out of five high-resolution uncorrected simulations outperform their 0.44° version (last column  
327 of Table 5). Similarly, for the Calder catchment, one 0.11° simulation outperforms its 0.44° version and all  
328 five high-resolution simulations outperform their low-resolution version for the Glaslyn and Coquet  
329 catchments. This indicates that topography has an influence in the simulation of temperature and RCM  
330 resolution has an effect in the simulation skill for catchments with larger elevation variability where, for  
331 observations at high elevation, the 0.44° RCMs would be expected to have positive biases as the grid  
332 elevation is lower than the observations.

333 Based on the rank, the overall best performing simulation for the upper Thames and Calder  
334 catchments is 0.44° RACMO, whereas for the Glaslyn and Coquet catchments, the 0.11° RACMO and  
335 HIRHAM simulations, respectively, outperform the rest. This implies that biases from the high-resolution  
336 simulations are smaller for the catchments with complex topography, which is better represented by the  
337 0.11° simulations. The biases are a consequence of systematic model biases in the elevation and a lack of

338 representation of the elevation variability. Nevertheless, for larger and flatter catchments the simulation skill  
339 from both resolutions is similar.

### 340 **3.2.1.2. Precipitation**

341 Now we assess the skill of the uncorrected RCMs to simulate precipitation. Overall, RCMs have  
342 biases when simulating extremes. For instance, the SDII ratio is underestimated in all catchments by the  
343 0.44° simulations (Figs. 4a, S1a and S2a), except for the Coquet (Fig. S3a). In all catchments the RX1day  
344 is overestimated by both resolutions between 24% and 93%. The R10 and R20 are underestimated at the  
345 Glaslyn catchment between -23 and -77 days and between -16 and -45 days, respectively (Fig. S1d).  
346 Similarly, in the Calder catchment R10 and R20 are underestimated by the 0.44° simulations between -5  
347 and -10 days and between -3 and -4 days, respectively (Fig. S2d). These results indicate that the  
348 uncorrected models can provide unrealistic simulations of extreme precipitation.

349 It is expected that the models simulate the precipitation mean better than the extremes. Even  
350 though the spread of the models includes the observed mean precipitation for most catchments, there are  
351 cases when this does not happen. The annual mean precipitation is underestimated by both resolutions in  
352 the Glaslyn catchment between -22% and -67% (Fig. S1c). This may be because the analysed RCMs do  
353 not correctly simulate convective precipitation. In the Calder catchment the 0.44° simulations underestimate  
354 the annual mean precipitation between -7% and -16% (Fig. S2c). This can be due to local precipitation not  
355 being correctly simulated by the coarse models. The absolute monthly mean precipitation bias for both  
356 resolutions varies between 7% and 67% in all study cases (Figs. 4c, S1c, S2c and S3c).

357 The simulated precipitation bias spread increases in all catchments as the percentile increases.  
358 The spread of the 0.11° simulations is larger than for the 0.44° simulations (Fig. 5, first row). In the upper  
359 Thames catchment, the 0.11° simulations reach their largest spread, -1 to 4 mm/day, above the 95<sup>th</sup>  
360 percentile whereas the largest spread of the 0.44° RCMs ranges from -1 to 1 mm/day (Fig. 5a). In the  
361 Glaslyn catchment, the bias spread deviates from the observations at the 50<sup>th</sup> percentile for the 0.44°  
362 simulations and at the 60<sup>th</sup> percentile for the 0.11° simulations (Fig. 5d). In the Calder catchment, the 0.11°  
363 simulations spread includes the no bias threshold for the whole distribution whereas the 0.44° simulations

364 spread deviates from that threshold at the 70<sup>th</sup> percentile (Fig. 5g). In the Coquet catchment, the spread  
365 from both resolutions includes the zero bias threshold for almost all percentiles (Fig. 5j).

366 The dry and wet spell biases are important for the simulation of river flow as this is influenced by  
367 the daily sequence of the wet/dry conditions. The absolute dry spell bias for both resolutions in all  
368 catchments range between 0.2 to 1.6 days, with a similar simulation skill in all catchments (Figs. 4b, S1b,  
369 S2b, S3b). Likewise, the absolute wet spell bias for both resolutions varies between 0.1 and 1.6 days in all  
370 catchments (Figs. 4b, S1b, S2b, S3b). Biases in the upper Thames for this measure are smaller, 0.2 to 0.6  
371 days (Fig. 4b), compared to the other catchments. These results do not show large simulation biases.  
372 Considering the time-series based indices, correlation coefficients are above 0.4 and below 0.8 in all  
373 catchments, showing differences between the daily observations and simulations (Figs. 4c, S1c, S2c, S3c).

374 Considering the ranking for all indices, only for the Glaslyn catchment do all the 0.11° simulations  
375 outperform their 0.44° version (Table 6, last column). From the five RCMs, two 0.11° simulations outperform  
376 their low-resolution version for the Upper Thames and three for the Calder and Coquet catchments. The  
377 0.11° CCLM and WRF have better simulation skill than their 0.44° version in all catchments. In contrast, for  
378 HIRHAM and RCA, the improvement is only observed in one catchment. For the latter models, there is no  
379 added value from increasing the resolution as the simulation processes occurring at higher resolutions than  
380 the 0.44° gridbox do not improve the results, possibly due to an inappropriate physical representation. The  
381 0.11° CCLM is the best performer in all catchments, except for the Glaslyn where 0.11° HIRHAM has the  
382 highest rank.

383 All high-resolution simulations outperform their coarse simulations at the Glaslyn catchment due to  
384 the differences between the sizes of the catchment and the different cells. Thus, increasing the RCM  
385 resolution increases their simulation skill for catchments with larger elevation variability because the RCMs  
386 are able to represent the high-resolution features. In general, increasing the RCM resolution reduces the  
387 simulation biases in the upper tail of the distribution, but there are also high-resolution models that  
388 consistently overestimate precipitation (e.g. RCA in Figs. 4, S1, S2, S3). The low-resolution models do not  
389 simulate the small-sized catchment accurately. In contrast, the flat and large catchments are simulated  
390 similarly by both resolutions, showing no added value from increasing RCM resolution.

### 391                    **3.2.1.3.    River flow**

392                    Now, we evaluate the RCM skill in providing inputs for simulating the river flow in each catchment.  
393                    In the upper Thames, the 0.11° RCMs overestimate the spring discharge by between 16% and 194% (Fig.  
394                    6a). Both resolutions underestimate all indices in the Glaslyn catchment (Fig, S4). In the Calder catchment,  
395                    the 0.44° RCMs underestimate the annual (-9% to -31%) and autumn (-10% to -50%) flows, whereas the  
396                    0.11° RCMs overestimate the discharge during winter (3% to 63%) and spring (22% to 104%) (Fig. S5a).  
397                    Also, the Q10 and Q10 annual frequency are underestimated by the 0.44° RCMs (Fig. S5b and c). In the  
398                    Coquet catchment, the winter mean discharge is underestimated by the 0.44° RCMs by between -7% and  
399                    -42% and during summer it is overestimated by the 0.11° RCMs by between 2% and 218% (Fig. S6a). In  
400                    addition, the Q95 is overestimated by the 0.11° simulations.

401                    Except for the Glaslyn catchment, the multi-model simulation spread of the flow duration curve  
402                    (FDC) from both resolutions includes the observed FDC entirely (Fig. 7, first row). For the Glaslyn  
403                    catchment, both resolutions underestimate the FDC with the 0.11° simulation spread being closer to the  
404                    observed FDC (Fig. 7d). The 0.44° simulation spread is larger than the 0.11° spread in the Coquet, but  
405                    smaller in the upper Thames. In the remaining catchments, the spreads of both resolutions are similar.

406                    Overall, the maximum monthly NSE values are 0.42 for the Upper Thames (Fig. 6e), 0.22 for the  
407                    Glaslyn (Fig. S4e), 0.67 for the Calder (Fig. S5e) 0.26 for the Coquet catchment (Fig. S6e), indicating that  
408                    the best river flow simulation is moderate to poor for all catchments except for the Calder. In contrast, the  
409                    minimum NSE values are negative in all catchments, implying that there are RCM outputs that generate  
410                    unreliable river flow simulations even at the monthly times step. Negative NSE values can be a result of  
411                    river flow overestimation in all indices, for instance 0.11° RCA and HIRHAM in the Calder and Coquet  
412                    catchments. The Spearman correlation coefficients of the daily river flow are higher for the upper Thames  
413                    and Calder and smaller for the Glaslyn and Coquet, indicating that the RCMs are able to simulate the daily  
414                    river flow sequence better on the large and flat sites compared to the small and topographically-complex  
415                    catchments (Fig. 67f, S4f, S5f and S6f).

416                    Comparing their skill in simulating all indices by means of their rank, three 0.11° simulations  
417                    outperform their 0.44° version in the Upper Thames, five in the Glaslyn, one in the Calder and two in the



418 Coquet catchment (Table 7, last column). Overall, for both resolutions, biases in particular indices are large  
419 and the skill of the pairwise indices (NSE, MSE, correlation) is low. The 0.11° simulation biases are  
420 consistently smaller than the 0.44° biases only for the Glaslyn catchment due to the difference between the  
421 catchment and the 0.44° RCM cell size. However, for this catchment biases are large even for the high-  
422 resolution simulations indicating that subgrid processes that result in precipitation increases are not  
423 represented by the models. Only CCLM gives better simulation skill for its high-resolution in all catchments.

### 424 **3.2.2. Bias-corrected RCM simulations**

#### 425 **3.2.2.1. Temperature**

426 Bias-correction reduces the mean and percentile biases by construction (Figs. 3e,f,g,h). Thus, the  
427 skill of all RCMs becomes similar in all catchments, as expected. Overall, the larger distribution biases are  
428 for the 1<sup>st</sup> and 99<sup>th</sup> temperature percentiles, with biases lower than 1°C (Figs. 2, i). Even though these  
429 percentiles have the largest biases after bias correction, as may be expected the biases are smaller than  
430 those of the uncorrected RCMs. QM does not improve the daily sequence simulation. As a consequence,  
431 there is only a slight change in the Pearson correlation coefficient of the daily time series (Figs. 2, iii).

#### 432 **3.2.2.2. Precipitation**

##### 433 **3.2.2.2.1. Gamma distribution QM**

434 The skill of both RCM resolutions becomes similar after application of GQM. Nevertheless, biases  
435 are not reduced for the 95<sup>th</sup> percentile, SDII ratio, wet spell length, R95p and R20 in the Upper Thames, for  
436 RX1day in the Calder and for the SDII ratio in the Coquet catchment. These indices evaluate the extremes,  
437 which are inflated by the correction method (Cannon et al., 2015), and the precipitation intensity.

438 Considering the indices that are not based on the distribution, the Spearman correlation slightly  
439 increases after GQM (Figs. 4c, S1c, S2c and S3c) whereas for the MSE the multi-model ensemble bias is  
440 reduced, but there are cases when the biases of individual RCMs increase (Figs. 4c, S1c, S2c and S3c).  
441 The same happens for the wet and dry spell lengths (Figs. 4b, S1b, S2b and S3b) and RX1day (Figs. 4c,  
442 S1c, S2c and S3c). The multi-model bias spread from both resolutions is similar and smaller than 1 mm/day  
443 up to the 90<sup>th</sup> percentile in all catchments (Fig. 5, second row). Above the 90<sup>th</sup> percentile, the spread of both

444 resolutions increases exponentially. The bias spread in the extremes is larger for the Glaslyn catchment  
445 possibly as a consequence of the bias magnitude of the original uncorrected simulation (Fig. 5e).

#### 446 **3.2.2.2. Double Gamma distribution QM**

447 After applying the DGQM method, the skill with respect to distribution-based indices from all RCMs  
448 at both resolutions becomes similar. The biases for most distribution-based indices are reduced compared  
449 to both uncorrected and GQM. In all catchments, the biases are lower than 1 mm/day below the 99<sup>th</sup>  
450 percentile after which biases increase. Thus, DGQM reduces the percentile biases in all catchments  
451 compared to GQM. For the 90<sup>th</sup> precipitation percentile the DGQM approach increases the biases in all  
452 catchments because at this percentile the method segments the precipitation distribution, generating an  
453 increment in the bias. Nevertheless, this increase is approximately  $\pm 1$  mm/ day in all catchments except  
454 the Glaslyn. Additionally, the simulation bias spread of both resolutions is similar for all catchments, as  
455 expected (Fig. 5, last row).

456 For the extreme and precipitation intensity measures, DGQM reduces the biases compared to GQM  
457 except for the RX1day and SDII ratio in the Upper Thames , R20 in the Glaslyn, R10 in the Calder and the  
458 SDII ratio in the Coquet catchment. The simulation skill of the uncorrected models and the GQM and DGQM  
459 approaches is similar in all catchments for the Spearman daily correlation coefficient. Overall, the DGQM  
460 provides outputs with smaller biases for most of the indices compared to the uncorrected and GQM  
461 simulations.

#### 462 **3.2.2.3. River flow**

##### 463 **3.2.2.3.1. Gamma distribution QM**

464 River flow is simulated using the GQM precipitation and temperature as drivers. GQM decreases  
465 the bias of all indices in every catchment, except for the Q10 in the upper Thames catchment (Fig. 6c). The  
466 bias-corrected FDC simulation spread decreases for both resolutions in all catchments (Fig. 7, second row).  
467 The observed FDC is completely included within the spread of both resolutions showing a good simulation  
468 of the entire distribution.

469 From the pairwise indices, the skill of the multi-model ensemble improves for the monthly NSE (Fig.  
470 6e) and the spread of the daily MSE is reduced in most cases. However, GQM can result in negative NSE

471 values for some models that had positive values when these were not bias-corrected (e.g. 0.44° RACMO  
472 and HIRHAM in the Upper Thames). The Spearman correlation of daily time series increases slightly in all  
473 cases (Fig. 6f, S4f, S5f and S6i).

#### 474 **3.2.2.3.2. Double Gamma distribution QM**

475 The DGQM approach decreases the biases for all the distribution-based indices compared to both  
476 uncorrected and GQM with the exception of Q95 for the Glaslyn catchment. Considering the non-  
477 distribution-based indices, the NSE and MSE are not improved for the Coquet catchment. Even though the  
478 biases are reduced, the simulation skill among all RCMs does not become similar for specific cases with  
479 indices involving the extremes and the pairwise simulation (e.g. the Q10 annual frequency, Q10 and NSE  
480 for the Upper Thames, Fig. 6b,c and e). Overall, the daily MSE and monthly NSE simulation skill improves  
481 compared to the GQM approach. Thus implying that the river flow simulation skill is better when using the  
482 DGQM. By construction of the bias correction method, the FDC simulation spread of both resolutions is  
483 similar in shape and amplitude (Fig. 7, bottom row). Compared to GQM, the DGQM simulation spread is  
484 further reduced.

485 The Spearman correlation coefficient of the daily river flow time series increases slightly with not a  
486 large difference compared to the GQM simulations. Overall, applying the DGQM approach results in smaller  
487 biases compared to the GQM, in specific for the simulations of extremes and the monthly sequence.

### 488 **4. Discussion**

489 Regarding our first research question, as to whether the relative performance of the high- resolution  
490 simulations is better than that of the lower-resolution simulations, the results show that the high-resolution  
491 RCMs consistently have a better simulation skill for climate and river flow only in the Glaslyn catchment.  
492 This is mainly because the size of this catchment is smaller than the 0.44° RCM cell, and it has a complex  
493 topography and high precipitation. As a consequence, the skill of the 0.44° simulations in reproducing the  
494 local physical features of this catchment is not good. For the other catchments, all of which are larger in  
495 size and with less complex topography and less precipitation, both resolutions have a similar performance.  
496 Similar results were obtained for the Upper Danube using HIRHAM at resolutions of 50 km x 50 km and 12  
497 km x 12 km (Dankers et al., 2007). Only the skill of CCLM improved when using the high-resolution version.

498 Kotlarsky et al. (2014) found that CCLM also gave good results when simulating the mean, seasonal and  
499 95<sup>th</sup> percentile of precipitation over the British Isles. In our study, the remaining RCMs did not improve their  
500 simulation skill, implying that the high-resolution versions of these models do not accurately represent  
501 processes occurring at higher resolutions.

502         The performance of the two RCM resolutions at simulating temperature was clearly linked to the  
503 topographic characteristics of the study catchments. In the upper Thames and Calder catchments, which  
504 have relatively flat topography, we found that there is no clear added value from the uncorrected high-  
505 resolution RCMs; however, in the topographically-complex Glaslyn and Coquet catchments, all 0.11°  
506 simulations outperformed their 0.44° version. These findings are similar to that of Onol et al. (2012) and  
507 Tolika et al. (2016) and it is likely that they can be attributed to the difference in elevation from the grid cells  
508 of the observations and models, and the lack of representation of the spatial variability. Increases in the  
509 simulation skill of local climate when using higher-resolution simulations have been reported before,  
510 particularly for mountainous regions (Evans et al., 2013; Larsen et al., 2013; Tolika et al., 2016).

511         The uncorrected 0.11° simulations largely underestimate the precipitation and river flow  
512 observations of the Glaslyn catchment, mainly due to the catchment's topographic complexity and high  
513 levels of precipitation. Similar results for the Euro-CORDEX RCMs have been obtained for precipitation in  
514 other regions with complex topography (e.g. Casanueva et al., 2016; Prein et al., 2015; Torma et al., 2015).  
515 For the remaining catchments, the multi-model simulation spread of the simulations of both resolutions  
516 includes the observed FDC, indicating that the models are able to provide useful simulations that resemble  
517 the observed river flow. However, the simulation spread can be large; deviations in the annual mean river  
518 flow reach almost 100% for some RCMs. Individual uncorrected RCMs have small biases and satisfactory  
519 simulations of the river flow (e.g. 0.11° CCLM in the Calder and Coquet catchments), but there are also  
520 RCMs that are not able to provide useful simulations. For example, the 0.11° RCA had the largest  
521 precipitation and river flow biases in most indices for all catchments. In contrast, all the bias-corrected RCM  
522 simulations are closer to the observed climate and river flow. Furthermore, the simulation skill from all bias-  
523 corrected RCMs at both resolutions becomes similar and as a result, the simulation spread of the multi-

524 model ensemble is reduced compared to the uncorrected simulations, providing a smaller range of possible  
525 scenarios.

526 Our results show that uncorrected RCMs provide river flow simulations that have too much spread  
527 to be able to be used for impact studies (also stated by Kay et al., 2015; Cloke et al., 2013). Both resolutions  
528 have a similar performance when simulating the seasonal mean river flow as there are biases from both  
529 resolutions. However, certain high-resolution models tend to overestimate the seasonal flow largely for  
530 most of the catchments and seasons (e.g. RCA in all catchments and HIRHAM in the Coquet and Calder  
531 catchments). In contrast, the low-resolution CCLM underestimates river flow for all seasons and  
532 catchments. For the medium-sized Calder catchment, individual models have different biases per season  
533 but the multi-model ensemble mean shows a consistent underestimation for high-resolution models and  
534 underestimation of river flow for the low-resolution modes. This is not distinguished in the larger Upper  
535 Thames nor in the Coquet catchment. Similar to the annual mean flow, both resolutions underestimate the  
536 seasonal flow in the Glaslyn catchment. In comparison, all the bias-corrected RCMs simulate the river flow  
537 much closer to the observed flows and reduce the simulation spread, thus providing plausible inputs for  
538 impact studies.

539 Finally, to answer our last research question, we evaluate the simulation skill of DGQM compared  
540 to GQM. Using four catchments with different characteristics, the DGQM provides a better simulation of the  
541 river flow characteristics compared to the QGM approach, with a higher improvement for the simulation of  
542 extremes and the monthly sequence. The GQM systematically reduces the precipitation bias up to the 90<sup>th</sup>  
543 percentile, but exponentially increases the bias above this percentile. Therefore, to capture the properties  
544 of extremes, we suggest using the DGQM with the 90<sup>th</sup> percentile as segmentation threshold in contrast to  
545 Yang et al. (2010) who divided the distribution at the 95<sup>th</sup> percentile. Based on our results, the DGQM  
546 reduces the precipitation and river flow biases of most indices compared to the commonly used GQM. This  
547 is particularly relevant for the analysis of extreme precipitation and high flows as the GQM is usually  
548 employed in flood analysis (e.g. Cloke et al., 2013) and river flow projections (e.g. Prudhomme et al., 2013).  
549 In addition, the DGQM reduces the ensemble spread more than the GQM, without introducing much extra  
550 complexity. However, no bias correction method will remove all biases. Thus, the selection of the method

551 depends on the requirements of each study (Nguyen et al., 2017) and it should be tested to evaluate  
552 whether the benefits justify their calculation complexities.

553 Ideally, RCMs should not require post-processing techniques to provide simulations which can be  
554 used with confidence (Ehret et al., 2012). However, our results demonstrate large biases for various  
555 diagnostic indices for the reanalysis-driven RCMs. Particular RCMs provide plausible river flow simulations,  
556 for instance, 0.11° CCLM for the Calder catchment when assessing the annual and seasonal means, low  
557 flows, high flow occurrence and pairwise indices. However, the RCM simulation skill is catchment-  
558 dependent. Thus, at the moment, bias correction seems to be the best approach to reduce the ensemble  
559 spread and its biases. Nevertheless, bias correction methods should be used carefully for the analysis of  
560 future projections (Cloke et al., 2013) as bias correction cannot correct fundamental problems from the  
561 original climate model (Maraun and Widmann, 2015; Maraun et al., 2017) and the spread of the bias-  
562 corrected simulations might not reflect the total real uncertainty. Climate research is focusing on  
563 determining the causes behind the biases (e.g. Addor et al., 2016) and improving the simulation of the  
564 processes (e.g. Zittis et al., 2017; Meredith et al., 2015). For instance, convection permitting models seek  
565 to improve the simulation of precipitation extremes (Tölle et al., 2017; Gutjahr et al., 2016). However, the  
566 computational cost of developing such models is large and, as a consequence, the simulation length is  
567 short and the availability of GCM-RCM projections is low.

568 By analysing four catchments with different characteristics, we evaluate the RCM simulation skill in  
569 different contexts. Our results suggest that the small size and the high precipitation (e.g. Glaslyn catchment)  
570 are the main factors related to the better simulation skill from the high-resolution RCMs over the low-  
571 resolution models for the simulation of river flow. The importance of topographical complexity and other  
572 characteristics for the simulation outputs is secondary. This is highlighted by the results of the medium-  
573 sized Coquet catchment, for which both resolutions have similar simulation skill even with its complex  
574 topography. Although the hydrological model used (HEC-HMS) was chosen as it has been used before in  
575 assessment of climate change impacts (e.g. Babel et al., 2014; Azmat et al., 2015) and acknowledging that  
576 there are a diversity of methods used to simulate the hydrological processes, we note that our results are  
577 unlikely to substantially change when using other hydrological model. We support our statement as we

578 assess the performance of the different resolutions by evaluating the RCM outputs as well as the  
579 hydrological model outputs, both giving similar results. An assessment of the hydrological model uncertainty  
580 is beyond the scope of this study, but will be the subject of future work.

## 581 **5. Conclusions**

582 This study provides information on the added value from increasing RCM resolution and bias  
583 correction techniques for the simulation for river flow. Previous studies have assessed the improvement in  
584 the simulation skill of climate variables due to an increase in the RCM resolution, but this might not  
585 guarantee an improvement in the simulation of the river flow parameters that are relevant for impact studies.  
586 We conducted a comprehensive analysis on how the uncorrected and bias-corrected RCM outputs drive  
587 the simulations of river flow at high and low resolutions. Each RCM used here has the same  
588 parameterisation, domain and driving data at both resolutions, and therefore the comparison only evaluates  
589 the effect of increasing its resolution. We analysed four catchments located at different latitudes within  
590 Great Britain. These catchments vary in climate (e.g. precipitation ranging from 2900 mm yr<sup>-1</sup> to 762 mm  
591 yr<sup>-1</sup>), physical characteristics (flat and complex topographies, areas ranging from 69 km<sup>2</sup> to 1616 km<sup>2</sup>), land  
592 use (varying from urban-dominant to agricultural and natural areas) and hydrological characteristics (e.g.  
593 annual mean river flow ranging from 15.3 m<sup>3</sup> s<sup>-1</sup> to 5.8 m<sup>3</sup> s<sup>-1</sup>). We applied a detailed assessment of the  
594 simulation skill of the climate and hydrological models using a set of indices relevant for the analysis of  
595 different impacts.

596 We show that the uncorrected 0.11° RCMs only showed better skill in simulating precipitation and  
597 river flow in the small catchment. This is because the spatial resolution of the 0.44° models is four-times  
598 larger than the catchment size, whereas one cell of the 0.11° model is similar in area to the catchment.  
599 Nevertheless, the high-resolution simulations are not able to accurately represent the complex topography  
600 of this catchment and do not resolve local processes, underestimating the observed precipitation and the  
601 entire FDC. In Australia, Parana Manage et al. (2016) also found that the averaging of topography of gridded  
602 outputs influences on the accurate simulation of rainfall.

603 Both resolutions capture the temperature and precipitation distribution, as well as the FDC, for the  
604 remaining sites. Thus, in principle, the simulations could be used for climate change assessments.

605 Nevertheless, for most of the indices, the multi-model variability is large (e.g. the mpe of the annual mean  
606 river flow simulation ranges from 198% to -31% in the Upper Thames, with an average of 49%), making  
607 any interpretation difficult in practice. Only one RCM (CCLM) improves the river flow simulation when using  
608 its high-resolution version in all catchments, implying that the remaining models do not simulate the relevant  
609 high-resolution processes accurately as there is no consistent difference between their high and low  
610 resolution versions. Therefore, there is no added value from using the high-resolution RCMs in those  
611 catchments for the assessment of river flow impacts.

612 Bias-correction reduces the distribution-based biases for all RCMs and resolutions by construction.  
613 Thus, the bias-corrected high- and low-resolution RCMs have similar simulation skill for the distribution-  
614 based indices. There is also less spread from the ensemble simulation of precipitation and river flows (e.g.  
615 the mpe of the annual mean river flow simulations for the Upper Thames ranges from -1% to 16% when  
616 corrected using DGQM, with an average of 7%). Nevertheless, daily pairwise indices, which assess the skill  
617 of the model when simulating the observed time series, are not improved by bias correction. However, the  
618 monthly NSE results indicate that bias correction can improve the pairwise simulation on monthly  
619 timescales. Overall, correcting the RCMs to the local temperature and precipitation provides a reduction of  
620 the ensemble spread, making the outputs more useful for the analysis of impacts. Nevertheless, it should  
621 be considered that the ensemble spread of uncorrected and corrected models can underestimate the true  
622 simulation uncertainty.

623 In comparison to GQM, DGQM provides a larger reduction in the simulation biases for precipitation  
624 and river flow. The main difference between both methods is the greater correction from DGQM for  
625 precipitation extremes (95th percentile, R10, R20, R95p) and high flows (Q10 and Q10 annual frequency).  
626 The monthly NSE consistently shows an improvement in the simulation skill of RCMs that are corrected  
627 using DGQM. Overall, for most of the RCMs and considering the results from all indices, the DGWM  
628 outperforms GQM.

629 Our study shows that an increase in RCM resolution does not always imply a better simulation of  
630 hydrological impacts, especially for large catchments. In contrast, small catchments with complex  
631 topography are still difficult to be simulated accurately by high-resolution models, concurring with Dankers



632 et al (2007). The uncorrected RCM ensemble generally provides a large spread which makes it difficult to  
633 use for impact assessment. Similar outcomes have been obtained for other regions, for example Australia  
634 (Lockart et al. 2016), Canada and China (Wang et al. 2019). Bias-correction provides an alternative to  
635 reduce the biases and multi-model spread, making decision-making easier. From the methods evaluated  
636 here, DGQM reduces most of the RCM biases without much more complexity added to the bias-correction  
637 method employed when using GQM. However, and agreeing with Cloke et al. (2013), Huang et al. (2014)  
638 and Lockart et al. (2016), the bias-corrected outputs should be used carefully when evaluating changes in  
639 very extreme flows as the correction inflates the simulated extremes. Compared to previous studies, we  
640 can state that our results are robust as we included a larger number of RCMs with different  
641 parameterisations for our analysis.

642 Whilst effective, bias correction adds extra uncertainty to the analysis chain (Cloke et al., 2013;  
643 Rummukainen, 2016). Therefore, it must be used with consideration of its limitations: dependence on the  
644 training period (Lafon et al., 2013), assumption of temporal stability of the correction function (Chen et al.,  
645 2015), issues of sub-grid variability and inflation of variance (Maraun, 2013), inter-variable consistency  
646 (Wilcke et al., 2013), spatial representation over complex terrain (Maraun and Widmann, 2015) and biases  
647 from the driving data (Maraun et al., 2017). The extent to which the climate change signal is altered must  
648 also be considered (Maraun, 2013; Velázquez et al., 2015) along with the possibility that bias correction  
649 can produce larger biases for extremes than for the mean (Huang et al., 2014). Additionally, we  
650 acknowledge that using different data to drive the RCMs used in this study, for instance a GCM, could give  
651 different results, as could the use of a different hydrological model.

652 Our results can provide an insight on whether RCMs of high(er) resolution improve the simulation skill.  
653 These can be useful for regions of similar characteristics where high(er)-resolution RCMs have not been  
654 developed yet and would require considerable time and effort to be produced. If used, bias-correction  
655 methods should be tested for the specific analysis that will be performed. This study provided different  
656 methods to perform this testing for the different RCMs and bias-correction methods for climatology and  
657 hydrology.

658

659 **Acknowledgements**

660 Pastén-Zapata received funding from Consejo Nacional de Ciencia y Tecnología (Conacyt) and Secretaría  
 661 de Educación Pública (SEP) during the development of this study. We also thank the Euro-CORDEX  
 662 initiative for providing the climate simulations and the US Army Corp of Engineers for making the HEC-  
 663 HMS model available. We are also grateful for the historical observational datasets provided by the UK's  
 664 CEH and NERC. Finally, we thank two anonymous reviewers for their comments regarding this paper.

665  
 666  
 667

668 **Tables**

669

Table 1. Characteristics of the study sites

	Upper Thames	Glaslyn	Calder	Coquet
Area (km <sup>2</sup> )	1616	69	316	346
Maximum elevation (masl <sup>1</sup> )	330	1080	556	775
Minimum elevation (masl <sup>1</sup> )	52	30	40	71
Mean annual precipitation (mm/year)	762	2957	1251	968
Mean annual temperature (°C)	9.7	8.1	8.4	7.4
Mean annual PET (mm/yr)	522	477	486	473
Mean annual river flow (m <sup>3</sup> /s)	15.3	5.8	8.8	6.1
Precipitation 90th percentile (mm/day)	6.7	24.4	10.3	7.7
Precipitation 95th percentile (mm/day)	10.2	34.2	14.8	11.9
<sup>2</sup> Q10 (m <sup>3</sup> /s)	34.8	13.5	19.9	12.4
<sup>3</sup> Q95 (m <sup>3</sup> /s)	1.90	0.55	1.99	0.84

670 <sup>1</sup> Meters above the sea level

671 <sup>2</sup> River flow that is exceeded for 10% of the daily river flow time series

672 <sup>3</sup> River flow that is exceeded for 95% of the daily river flow time series

673

674

675

Table 2. RCMs used in this study

RCM	Institute	Period	Reference
CCLM-CLMCOM	Brandenburg University of Technology (BTU)	1989-2008	Böhm et al., 2006; Rockel et al., 2008
HIRHAM 5	Danish Meteorological Institute (DMI)	1989-2008	Christensen et al., 1998
RACMO22E	Royal Netherlands Meteorological Institute (KNMI)	1979-2008	Van Meijgaard et al., 2012
RCA4	Swedish Meteorological and Hydrological Institute (SMHI)	1984-2008	Samuelsson et al., 2011
WRF 3.3.1	Institute Pierre Simon Laplace (IPSL) and Institute National de	1989-2008	Skamarock et al., 2008

676

	l'Environnement Industriel et des Risques (INERIS)		
--	--	--	--

Table 3. Description of the precipitation, temperature and river flow indices used in this study

Index	Description	Performance measure
<b>Precipitation</b>		
95 <sup>th</sup> percentile	A measure of very extreme events: 95 <sup>th</sup> percentile of daily precipitation	Bias (mm/day)
90 <sup>th</sup> percentile	A measure of extreme events: 90 <sup>th</sup> percentile of daily precipitation	Bias (mm/day)
50 <sup>th</sup> percentile	50 <sup>th</sup> percentile of daily precipitation	Bias (mm/day)
25 <sup>th</sup> percentile	25 <sup>th</sup> percentile of daily precipitation	Bias (mm/day)
<sup>a</sup> Wet spell length	Mean wet spell length for a given month of the year	Bias (days)
<sup>a</sup> Dry spell length	Mean dry spell length for a given month of the year	Bias (days)
<sup>a</sup> Annual mean precipitation	Annual accumulated precipitation	Mean percentage error
<sup>a</sup> Monthly mean precipitation	Accumulated precipitation for a given month of the year	Mean percentage error
<sup>b</sup> Relative daily MSE	Mean daily square error, shown as ratio to the largest MSE result (considering both corrected and uncorrected RCMS)	MSE (ratio)
<sup>b</sup> Spearman correlation coefficient	Spearman correlation coefficients between the daily simulated and observed time series	Index
<sup>a</sup> Maximum one day precipitation (RX1day)	Maximum one-day precipitation for a given month of the year	Mean percentage error
<sup>a</sup> Simple Daily Intensity Index (SDII)	Ratio of the annual total precipitation to the number of wet days ( $\geq 1$ mm) in all years	Index
<sup>a</sup> Number of heavy precipitation days (R10)	Mean number of days with precipitation $\geq 10$ mm within a year	Bias (days)
<sup>a</sup> Number of very heavy precipitation days (R20)	Mean number of days with precipitation $\geq 20$ mm within a year	Bias (days)
<sup>a</sup> Very wet days (R95p)	Mean annual accumulated precipitation from days > 95th percentile in all years	Mean percentage error
<b>Temperature</b>		
<sup>a</sup> Annual mean temperature	Annual mean temperature over the validation period	Mean percentage error
<sup>a</sup> Monthly mean temperature	Monthly mean temperature	Mean percentage error
99 <sup>th</sup> percentile of daily mean temperature	99 <sup>th</sup> percentile of the daily mean temperature	Bias ( $^{\circ}$ C/day)
1 <sup>st</sup> percentile of daily mean temperature	1 <sup>st</sup> percentile of the daily mean temperature	Bias ( $^{\circ}$ C/day)
<sup>b</sup> Pearson correlation coefficient	Pearson correlation coefficient between the daily RCM and observation time series	Index
<b>River Flow</b>		
Q10	A measure of high flows: river flow that is exceeded for 10% of the daily river flow time series	Bias (m <sup>3</sup> /s)
Q95	A measure of low flows: river flow that is exceeded for 95% of the daily river flow time series	Bias (m <sup>3</sup> /s)
<sup>a</sup> Annual Q10 frequency	Mean number of days for which the observed Q10 is exceeded within a year	Bias (days)
<sup>a</sup> Annual mean river flow	Annual mean daily river flow over the validation period	Mean percentage error
<sup>a</sup> Winter (DJF) mean river flow	Winter mean daily river flow over the validation period	Mean percentage error
<sup>a</sup> Spring (MAM) mean river flow	Spring mean daily river flow over the validation period	Mean percentage error
<sup>a</sup> Summer (JJA) mean river flow	Summer mean daily river flow over the validation period	Mean percentage error
<sup>a</sup> Autumn (SON) mean river flow	Autumn mean daily river flow over the validation period	Mean percentage error
<sup>b</sup> Monthly NSE	Monthly Nash Sutcliffe Efficiency index	Index
<sup>b</sup> Relative daily MSE	Mean daily square error, shown as ratio to the largest MSE result (considering both corrected and uncorrected RCMS)	MSE (ratio)
<sup>b</sup> Spearman correlation coefficient	Spearman correlation coefficient between the daily simulated and observed time series	Index

<sup>a</sup> Estimated using the long term mean (one value over the entire series)

<sup>b</sup> Estimated considering the time series values (one value per time step)

679  
680

Table 4. Indices from the calibration and validation of the hydrological models

Catchment	Step	Period	Daily NSE	Q10 bias		Q95 bias	
				(m <sup>3</sup> /s)	(%)	(m <sup>3</sup> /s)	(%)
Upper Thames	Calibration	1986-2010	0.70	-2.1	-6	-0.45	-25
	Validation	1961-1985	0.57	1.5	5	-0.90	-44
Glaslyn	Calibration	1991-2010	0.78	1.0	8	-0.07	-11
	Validation	1971-1990	0.78	0.7	5	-0.03	-6
Calder	Calibration	1994-2010	0.62	1.5	8	-0.31	-16
	Validation	1976-1993	0.60	1.3	7	-0.24	-12
Coquet	Calibration	1992-2010	0.63	1.3	11	-0.24	-27
	Validation	1973-1991	0.52	-0.6	-5	-0.25	-31

681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706

707  
708

Table 5. RCM rank for the temperature indices for each catchment: 1 = best, 10 = worst. The asterisks (\*) indicate the resolution with the best simulation skill of each RCM in each catchment

		99th percentile	1st percentile	Annual mean	Monthly mean	Correlation	Average score	Ranking	
Upper Thames	0.11°CCLM	10	7	2	9	1	5.8	6	*
	0.11°HIRHAM	3	9	3	5	6	5.2	5	
	0.11°RACMO	2	8	9	7	4	6.0	7	
	0.11°RCA	7	5	10	10	5	7.4	10	
	0.11°WRF	4	1	5	4	8	4.4	2	*
	0.44°CCLM	9	10	1	8	2	6.0	7	
	0.44°HIRHAM	1	6	4	3	9	4.6	3	*
	0.44°RACMO	5	4	7	2	3	4.2	1	*
	0.44°RCA	8	2	6	1	7	4.8	4	*
0.44°WRF	6	3	8	6	10	6.6	9		
Glaslyn	0.11°CCLM	9	2	4	3	1	3.8	3	*
	0.11°HIRHAM	7	6	2	4	7	5.2	5	*
	0.11°RACMO	3	7	1	1	4	3.2	1	*
	0.11°RCA	2	4	3	2	6	3.4	2	*
	0.11°WRF	4	8	5	6	10	6.6	7	*
	0.44°CCLM	10	1	6	5	2	4.8	4	
	0.44°HIRHAM	8	3	8	7	9	7.0	8	
	0.44°RACMO	5	5	7	8	3	5.6	6	
	0.44°RCA	6	9	9	9	5	7.6	9	
0.44°WRF	1	10	10	10	8	7.8	10		
Calder	0.11°CCLM	9	7	8	8	1	6.6	7	
	0.11°HIRHAM	5	9	7	7	5	6.6	7	
	0.11°RACMO	8	10	10	10	4	8.4	9	
	0.11°RCA	10	8	9	9	6	8.4	9	
	0.11°WRF	7	3	1	4	8	4.6	4	*
	0.44°CCLM	6	6	6	5	2	5	5	*
	0.44°HIRHAM	4	2	2	1	9	3.6	2	*
	0.44°RACMO	2	4	5	2	3	3.2	1	*
	0.44°RCA	3	1	4	3	7	3.6	2	*
0.44°WRF	1	5	3	6	10	5	5		
Coquet	0.11°CCLM	9	2	2	3	2	3.6	3	*
	0.11°HIRHAM	1	3	3	2	5	2.8	1	*
	0.11°RACMO	3	7	9	7	4	6.0	5	*
	0.11°RCA	7	6	8	4	6	6.2	6	*
	0.11°WRF	5	1	1	1	8	3.2	2	*
	0.44°CCLM	4	4	7	5	1	4.2	4	
	0.44°HIRHAM	10	8	5	6	9	7.6	9	
	0.44°RACMO	6	9	6	9	3	6.6	8	
	0.44°RCA	2	5	10	8	7	6.4	7	
0.44°WRF	8	10	4	10	10	8.4	10		

709

710 Table 6. RCM rank for precipitation: 1-best, 10-worst. The asterisks (\*) indicate the resolution with the best simulation skill of each RCM in each catchment

		Pr 95th	Pr 90th	Pr 50th	Pr 25th	Annual Mean	Monthly MSE	Dry Spell	Wet Spell	Monthly Mean	Correl.	SDII	R10	R20	R95p	RX1day	Average score	Ranking	
Upper Thames	0.11°CCLM	8	5	1	2	5	2	1	4	5	1	6	8	4	8	2	4.1	1	*
	0.11°HIRHAM	7	4	4	3	3	5	6	6	1	3	5	7	6	3	7	4.7	4	
	0.11°RACMO	3	2	9	8	7	3	4	5	4	10	9	3	5	5	9	5.7	8	
	0.11°RCA	10	10	10	10	10	10	10	10	10	6	2	10	10	10	8	9.1	10	
	0.11°WRF	1	1	6	7	6	8	7	3	8	5	7	1	3	1	3	4.5	2	*
	0.44°CCLM	9	9	2	1	8	4	3	8	7	2	4	9	8	9	1	5.6	7	*
	0.44°HIRHAM	5	6	3	5	2	7	5	9	3	7	3	4	1	4	5	4.6	3	*
	0.44°RACMO	4	3	5	6	4	1	2	2	2	9	8	5	7	6	6	4.7	4	*
	0.44°RCA	2	8	8	4	9	9	9	1	9	4	1	2	1	2	4	4.9	6	*
0.44°WRF	6	7	7	9	1	6	8	7	6	8	10	6	9	7	10	7.1	9	*	
Glaslyn	0.11°CCLM	5	5	8	2	5	5	6	5	5	1	5	5	5	5	5	4.8	5	*
	0.11°HIRHAM	1	1	6	5	1	3	5	3	2	3	1	3	1	1	1	2.5	1	*
	0.11°RACMO	3	3	3	9	3	1	3	8	3	2	3	2	3	2	4	3.5	3	*
	0.11°RCA	2	2	2	10	2	2	8	6	1	6	2	1	2	3	2	3.4	2	*
	0.11°WRF	4	4	1	6	4	4	4	4	4	7	4	4	4	4	3	4.1	4	*
	0.44°CCLM	10	9	10	3	9	9	9	9	9	5	9	9	9	8	7	8.3	9	
	0.44°HIRHAM	9	10	9	1	10	10	10	10	10	9	7	10	10	9	9	8.9	10	
	0.44°RACMO	7	7	4	7	7	7	2	1	7	4	10	7	7	7	8	6.1	7	
	0.44°RCA	8	8	7	4	8	8	7	7	8	8	8	8	8	10	10	7.8	8	
0.44°WRF	6	6	5	8	6	6	1	2	6	10	6	6	6	6	6	5.7	6		
Calder	0.11°CCLM	1	2	2	1	1	1	2	3	1	1	1	2	2	2	9	2.1	1	*
	0.11°HIRHAM	10	10	8	5	9	9	7	8	9	5	7	9	10	10	10	8.4	9	
	0.11°RACMO	2	1	9	9	3	5	5	9	4	4	4	1	1	1	3	4.1	2	*
	0.11°RCA	9	9	10	10	10	10	10	10	10	6	3	10	9	9	1	8.4	9	
	0.11°WRF	3	3	6	4	6	8	6	5	7	8	2	4	3	3	8	5.1	5	*
	0.44°CCLM	6	7	4	2	8	3	4	4	8	2	5	7	4	6	2	4.8	3	
	0.44°HIRHAM	4	4	1	3	7	4	9	6	6	7	6	3	5	4	5	4.9	4	*
	0.44°RACMO	8	8	7	7	5	2	3	1	3	3	10	8	8	8	6	5.8	7	
	0.44°RCA	7	6	3	6	4	7	8	2	5	9	8	6	7	7	7	6.1	8	*
0.44°WRF	5	5	5	8	2	6	1	7	2	10	9	5	6	5	4	5.3	6		
Coquet	0.11°CCLM	4	5	1	1	2	1	1	3	1	1	3	4	2	1	2	2.1	1	*
	0.11°HIRHAM	6	9	9	7	9	9	9	7	9	5	1	7	5	6	4	6.8	8	
	0.11°RACMO	5	3	6	8	1	3	7	5	2	4	9	5	4	5	5	4.8	4	*
	0.11°RCA	10	10	10	10	10	10	10	10	10	9	7	10	9	10	1	9.1	10	
	0.11°WRF	2	1	5	3	3	6	5	6	3	7	5	2	1	2	3	3.6	2	*
	0.44°CCLM	7	6	4	2	8	4	3	8	7	2	3	6	7	7	6	5.3	6	
	0.44°HIRHAM	3	2	8	9	4	5	8	2	4	8	1	1	3	4	8	4.7	3	*
	0.44°RACMO	8	7	3	4	6	2	2	4	6	3	9	8	10	9	10	6.1	7	
	0.44°RCA	1	4	7	5	5	8	6	1	5	6	7	3	6	3	7	4.9	5	*
0.44°WRF	9	8	2	6	7	7	4	9	8	10	5	9	8	8	9	7.3	9		

711 Table 7. RCM rank for river flow: 1-best, 10-worst. The asterisks (\*) indicate the resolution with the best simulation skill of each RCM in each catchment

		Annual mean	Winter mean	Spring mean	Summer mean	Autumn mean	Monthly NSE	Daily MSE	Correl.	Q10	Q10 frequency	Q95	Average score	Rank	
Upper Thames	0.11° CCLM	1	2	2	3	3	1	1	4	1	1	1	1.8	1	*
	0.11° HIRHAM	3	4	4	2	1	3	3	1	3	3	4	2.8	2	*
	0.11° RACMO	8	9	8	6	9	9	8	7	8	8	9	8.1	9	
	0.11° RCA	10	10	10	10	10	10	10	10	10	10	10	10.0	10	
	0.11° WRF	7	1	6	9	5	6	6	8	5	6	8	6.1	6	*
	0.44° CCLM	4	7	1	4	7	2	2	3	6	4	3	3.9	4	
	0.44° HIRHAM	2	5	3	1	2	4	5	5	2	2	2	3.0	3	
	0.44° RACMO	5	6	5	5	6	5	4	2	4	5	7	4.9	5	*
	0.44° RCA	9	8	9	7	8	8	9	6	9	9	6	8.0	8	*
0.44° WRF	6	3	7	8	4	7	7	9	7	7	5	6.4	7		
Glaslyn	0.11° CCLM	5	5	5	6	6	5	5	4	5	5	8	5.36	5	*
	0.11° HIRHAM	1	1	1	4	1	1	3	2	1	1	4	1.82	1	*
	0.11° RACMO	2	2	2	3	3	2	1	1	2	2	2	2	2	*
	0.11° RCA	3	3	3	1	2	3	2	6	3	3	1	2.73	3	*
	0.11° WRF	4	4	4	2	4	4	4	5	4	4	3	3.82	4	*
	0.44° CCLM	9	9	10	10	10	9	9	8	10	8	9	9.18	9	
	0.44° HIRHAM	10	10	9	9	9	10	10	9	9	8	10	9.36	10	
	0.44° RACMO	7	6	7	8	7	7	7	3	7	7	6	6.55	7	
	0.44° RCA	8	8	8	7	8	8	8	10	8	10	7	8.18	8	
0.44° WRF	6	7	6	5	5	6	6	7	6	6	5	5.91	6		
Calder	0.11° CCLM	2	2	6	4	5	2	1	2	1	1	1	2.45	2	*
	0.11° HIRHAM	9	10	9	8	9	9	9	4	9	9	9	8.55	9	
	0.11° RACMO	6	6	7	6	6	5	6	3	6	8	8	6.09	6	
	0.11° RCA	10	9	10	10	10	10	10	9	10	10	10	9.82	10	
	0.11° WRF	7	8	8	9	2	8	8	6	7	7	7	7	8	
	0.44° CCLM	8	5	5	7	8	6	3	5	8	6	6	6.09	6	
	0.44° HIRHAM	5	4	3	5	7	4	5	10	5	5	4	5.18	5	*
	0.44° RACMO	3	3	2	1	1	1	2	1	3	2	2	1.91	1	*
	0.44° RCA	4	7	4	3	3	7	7	8	4	4	5	5.09	4	*
0.44° WRF	1	1	1	2	4	3	4	7	2	3	3	2.82	3	*	
Coquet	0.11° CCLM	1	5	1	1	6	4	5	1	4	1	1	2.73	1	*
	0.11° HIRHAM	9	6	9	9	8	9	9	5	8	9	9	8.18	9	
	0.11° RACMO	7	1	7	7	5	6	3	6	2	2	8	4.91	5	
	0.11° RCA	10	10	10	10	10	10	10	10	10	10	10	10	10	
	0.11° WRF	2	2	2	5	2	2	6	4	1	3	5	3.09	2	*
	0.44° CCLM	8	9	8	4	9	7	2	2	9	8	2	6.18	7	
	0.44° HIRHAM	6	3	4	8	7	8	8	9	5	7	7	6.55	8	*
	0.44° RACMO	3	7	3	2	3	1	1	3	6	6	4	3.55	3	*
	0.44° RCA	5	4	6	6	4	5	7	7	3	4	6	5.18	6	*
0.44° WRF	4	8	5	3	1	3	4	8	7	5	3	4.64	4		



713           **References**

- 714    Addor, N., Rohrer, M., Furrer, R., Seibert, J., 2016. Propagation of biases in climate models from synoptic to the  
715            regional scale: Implications for bias adjustment. *J. Geophys. Res. Atmos.*, 121.  
716            doi:10.1002/2015JD024040.
- 717    Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration-Guidelines for computing crop  
718            water requirements-FAO Irrigation and drainage paper 56. FAO, Rome, 300(9), D05109.
- 719    Arnell, N.W. 2011. Uncertainty in the relationship between climate forcing and hydrological response in UK  
720            catchments. *Hydrology and Earth System Sciences*, 15(3), 897-912
- 721    Azmat, M., Laio, F., Poggi, D., 2015. Estimation of water resources availability and mini-hydro productivity in  
722            high-altitude scarcely-gauged watershed. *Water Resour Manage*, 29:5037-5054. doi: 10.1007/s11269-  
723            015-1102-z.
- 724    Babel, M.S., Bhusal, S.P., Wahid, S.M., Agarwal, A., 2014. Climate change and water resources in the Bagmati  
725            River Basin, Nepal. *Theor. Appl. Climatol.* 115,639-654. doi: 10.1007/s00704-013-0910-4.
- 726    Benestad, R.E., 2010. Downscaling precipitation extremes. *Theoretical and Applied Climatology*, 100(1-2), 1-21.
- 727    Böhm, U., Kücken, M., Ahrens, W., Block, A., Hauffe, D., Keuler, K., Rockel, B., Will, A., 2006. CLM-the climate  
728            version of LM: Brief description and long-term applications. *Tech. Rep.*, COSMO Newsletter.
- 729    Bucchignani, E., Montesarchio, M., Zollo, A.L. and Mercogliano, P., 2016. High-resolution climate simulations  
730            with COSMO-CLM over Italy: performance evaluation and climate projections for the 21st  
731            century. *International Journal of Climatology*, 36(2), pp.735-756.
- 732    Cannon, A.J., Sobie, S.R., Murdock, T.Q., 2015. Bias correction of GCM precipitation by quantile mapping: How  
733            well do methods preserve changes in quantiles and extremes? *Journal of Climate*, 28(17), 6938-6959.
- 734    Casanueva, A., Kotlarski, S., Herrera, S., Fernández, J., Gutiérrez, J. M., Boberg, F., Colette, A., Christensen,  
735            O. B., Goergen, K., Jacob, D., Keuler, K., Nikulin, G., Teichmann, C., Vautard, R., 2016. Daily  
736            precipitation statistics in a EURO-CORDEX RCM ensemble: added value of raw and bias-corrected  
737            high-resolution simulations. *Clim. Dyn.* 1-19. doi: 10.1007/s00382-015-2865-x

738 Chen, J., Brissette, F.P. and Lucas-Picher, P., 2015. Assessing the limits of bias-correcting climate model  
739 outputs for climate change impact studies. *Journal of Geophysical Research: Atmospheres*, 120(3),  
740 1123-1136.

741 Christensen, J.H., Boberg, F., Christensen, O.B., Lucas-Picher, P., 2008. On the need for bias correction of  
742 regional climate change projections of temperature and precipitation. *Geophys. Res. Lett.* 35(20),  
743 L20709. doi: 10.1029/2008GL035694.

744 Christensen, O.B., Christensen, J.H., MACHENHAUER, B., Botzet, M., 1998. Very high-resolution regional climate  
745 simulations over Scandinavia-present climate. *J. Clim.* 11(12), 3204-3229.

746 Christerson, B.V., Vidal, J.P., Wade, S.D. 2012. Using UKCP09 probabilistic climate information for UK water  
747 resource planning. *J. Hydrol.* 424-425: 48-67. doi: 10.1016/j.hydrol.2011.12.020.

748 Cloke, H.L., Wetterhall, F., He, Y., Freer, J.E., Pappenberger, F., 2013. Modelling climate impact on floods with  
749 ensemble climate projections. *Q. J. R. Meteorol. Soc.* 139, 282-297. doi: 10.1002/qj.1998.

750 Dankers, r., Christensen, O.B., Feyen, L., Kalas, M., de Roo, A., 2007. Evaluation of very high-resolution climate  
751 model data for simulating flood hazards in the Upper Danube Basin. *J. Hydrol.*, 347: 319-331. doi:  
752 10.1016/j.hydrol.2007.09.055.

753 Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A.,  
754 Balsamo, G., Bauer, P., Bechtold, P., 2011. The ERA-Interim reanalysis: Configuration and performance  
755 of the data assimilation system. *Quarterly Journal of the royal meteorological society*, 137(656), 553-  
756 597.

757 Di Luca, A., de Elía, R., Laprise, R., 2015. Challenges in the quest for added value of regional climate dynamical  
758 downscaling. *Curr. Clim. Change Rep.* 1, 10-21. doi: 10.1007/s40641-015-0003-9.

759 Eden, J.M., Widmann, M., Maraun, D., Vrac, M., 2014. Comparison of GCM- and RCM-simulated precipitation  
760 following stochastic postprocessing. *J. Geophys. Res. Atmos.* 119, 11040-11053.  
761 doi:10.1002/2014JD021732.

762 Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., Liebert, J., 2012. Should we apply bias correction to global  
763 and regional climate model data? *Hydrol. Earth Syst. Sci.* 16, 3391-3404. doi: 10.5194/hess-16-3391-  
764 2012.

765 Evans, J.P., McCabe, M.F., 2013. Effect of model resolution on a regional climate model simulation over  
766 southeast Australia. *Clim. Res.* 56:131-145. doi: 10.1007/s00382-012-1415-z.

767 Fantini, A., Raffaele, F., Torma, C., Bacer, S., Coppola, E., Giorgi, F., Ahrens, B., Dubois, C., Sanchez, E. and  
768 Verdecchia, M., 2016. Assessment of multiple daily precipitation statistics in ERA-Interim driven Med-  
769 CORDEX and EURO-CORDEX experiments against high resolution observations. *Climate Dynamics*, 1-  
770 24.

771 Feser, F., Rockel, B., von Storch, H., Winterfeldt, J., Zahn, M., 2011. Regional climate models add value to  
772 global model data. A review and selected examples. *Bull. Amer. Meteor. Soc.* September. doi:  
773 10.1175/2011BAMS3061.1.

774 Fowler, H.J. and Kilsby, C.G., 2007. Using regional climate model data to simulate historical and future river flows  
775 in northwest England. *Clim. Change.* 80: 337-367.

776 Gutjahr, O., Schefczyk, L., Reiter, P. and Heinemann, G., 2016. Impact of the horizontal resolution on the  
777 simulation of extremes in COSMO-CLM. *Meteorol. Z.*

778 Hough, M.N. and Jones, R.J.A., 1997. The United Kingdom Meteorological Office rainfall and evaporation  
779 calculation system: MORECS version 2.0-an overview. *Hydrology and Earth System Sciences*  
780 *Discussions*, 1(2), pp.227-239.

781 Huang, S., Krysanove, V., Hatterman, F. F., 2014. Does bias correction increase reliability of flood projections  
782 under climate change? A case study of large rivers in Germany. *Int. J. Climatol.* 34, 3780-3800. doi:  
783 10.1002/joc.3945.

784 Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O.B., Bouwer, L.M., Braun, A., Colette, A., Déqué, M.,  
785 Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N.,  
786 Jones, C., Keuler, K., Kovats, S., Kröner, N., Kotlarski, S., Kriegsman, A., Martin, E., van Meijgaard, E.,  
787 Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M.,  
788 Samuelsson, P., Somot, S., Soussana, J.F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., Yiou,  
789 P., 2014. EURO-CORDEX: new high-resolution climate change projections for European impact  
790 research. *Reg. Environ. Change.* 14,563-578. doi: 10.1007/s10113-013-0499-2.

791 Kay, A.L. and Davies, H.N., 2008. Calculating potential evaporation from climate model data: A source of  
792 uncertainty for hydrological climate change impacts. *Journal of Hydrology*, 358(3), 221-239.

793 Kay, A.L., Davies, H.N., Bell, V.A., Jones, R.G., 2009. Comparison of uncertainty for climate change impacts:  
794 flood frequency in England. *Clim. Change*. 92: 41-63. doi: 10.1007/s10584-008-9471-4.

795 Kay, A.L., Bell, V.A., Blyth, E. M., Crooks, S. M., Davies, H.N., Reynards, N. S. 2013. A hydrological perspective  
796 on evaporation: historical trends and future projections in Britain. *Journal of Water and Climate Change*,  
797 4(3) 193-208. doi: 10.2166/wcc.2013.014.

798 Kay, A.L., Jones, D.A., 2012. Transient changes in flood frequency and timing in Britain under potential  
799 projections of climate change. *Int. J. Climatol.* 32: 489-502. doi: 10.1002/joc.2288.

800 Kay, A.L., Rudd, A.C., Davies, H.N., Kendon, E.J., Jones, R.G., 2015. Use of very high resolution climate model  
801 data for hydrological modelling: baseline performance and future flood changes. *Clim. Change*. 133:193-  
802 208. doi: 10.1007/s10584-015-1455-6.

803 Keller, V.D.J., Tanguy, M., Prosdocimi, I., Terry, J.A., Hitt, O., Cole, S.J., Fry, M., Morris, D.G., Dixon, H., 2015.  
804 CEH-GEAR: 1 km resolution daily and monthly areal rainfall estimates for the UK for hydrological and  
805 other applications. *Earth System Science Data*, 7(1), 143-155.

806 Kingston, D.G., Todd, M.C., Taylor, R.G., Thompson, J.R., Arnell, N.W., 2009. Uncertainty in the estimation of  
807 potential evapotranspiration under climate change. *Geophysical Research Letters*, 36(20).

808 Kim, K.B., Kwon, H.H., Han, D., 2016. Precipitation ensembles conforming to natural variations derived from a  
809 regional climate model using a new bias correction scheme. *Hydrology and Earth System  
810 Sciences*,20(5), 2019-2034

811 Kotlarski, S., Keuler, K., Christensen, O.B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D.,  
812 van Meijgaard, E., Nikulin, G., Schär, C., Teichutard, R., Warrach-Sagi, K., Wulfmeyer, V., 2014.  
813 Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM  
814 ensemble. *Geosci. Model Dev.* 7, 1297-1333. doi: 10.5194/gmd-7-1297-2014.

815 Lafon, T., Dadson, S., Buys, G., Prudhomme, C., 2013. Bias correction of daily precipitation simulated by a  
816 regional climate model: a comparison of methods. *Int. J. Climatol.* 33:1367-1381. doi:10.1002/joc.3518.

817 Larsen, M.A., Thejll, P., Christensen, J.H., Refsgaard, J.C., Jensen, K.H., 2013. On the role of domain size and  
818 resolution in the simulations with the HIRHAM region climate model. *Climate dynamics*, 40(11-12), 2903-  
819 2918.

820 Lockart, N., Willgoose, G., Kuczera, G., Kiem, A.S., Chowdhury, A.K., Manage, N.P., Zhang, L. and Twomey, C.,  
821 2016. Case study on the use of dynamically downscaled climate model data for assessing water security  
822 in the Lower Hunter region of the eastern seaboard of Australia. *Journal of Southern Hemisphere Earth  
823 Systems Science*, 66(2), pp.177-202.

824 Macadam I., Argüeso, D., Evans, J.P., Liu, D.L., Pitman, A.J., 2016. The effect of bias correction and climate  
825 model resolution on wheat simulations forced with a regional climate model ensemble. *Int. J. Climatol.*  
826 doi: 10.1002/joc.4653.

827 Maraun, D., 2013. Bias correction, quantile mapping, and downscaling: revisiting the inflation issue. *J. Climate*.  
828 26, 2137-2143. doi: 10.1175/JCLI-D-12-00821.1.

829 Maraun, D., Shepherd, T.G., Widmann, M., Zappa, G., Walton, D., Gutiérrez, J.M., Hagemann, S., Richter, I.,  
830 Soares, P.M., Hall, A. and Mearns, L.O., 2017. Towards process-informed bias correction of climate  
831 change simulations. *Nature Climate Change*, 7(11), 3418.

832 Maraun, D., Wetterhall, F., Ireson, A.M., Chandler, R.E., Kendon, E.J., Widmann, M., Brienen, S., Rust, H.W.,  
833 Sauter, T., Themeßl, M., Venema, V.K.C., Chun, K.P., Goodess, C.M., Jones, R.G., Onof, C., Vrac, M.,  
834 Thiele-Eich., 2010. Precipitation downscaling under climate change: recent developments to bridge the  
835 gap between dynamical models and the end user. *Rev. Geophys.* 48, RG3003.  
836 doi:10.1029/2009RG000314.

837 Maraun, D., Widmann, M., 2015. The representation of location by a regional climate model in complex terrain.  
838 *Hydrol. Earth Syst. Sci.* 19, 3449-3456. doi: 10.5194/hess-19-3449-2015.

839 Maraun, D., & Widmann, M., 2018. *Statistical downscaling and bias correction for climate research*. Cambridge  
840 University Press.

841 Maraun, D., Widmann, M., Gutiérrez, J.M., Kotlarski, S., Chandler, E., Hertig, E., Wibig, J., Huth, R., Wilcke,  
842 R.A.I., 2015. VALUE: A framework to validate downscaling approaches for climate change studies.  
843 *Earth's Future*. 3, 1-14. doi: 10.1002/2014EF000259.

844 Maurer, E.P., Pierce, D.W., 2014. Bias correction can modify climate model simulated precipitation changes  
845 without adverse effect on the ensemble mean. *Hydrol. Earth Syst. Sci.* 18, 915-925. doi: 10.5194/hess-  
846 18-915-2014.

847 Mendoza, P.A., Mizukami, N., Ikeda, K., Clark, M.P., Gutmann, E.D., Arnold, J.R., Brekke, L.D. and  
848 Rajagopalan, B., 2016. Effects of different regional climate model resolution and forcing scales on  
849 projected hydrologic changes. *Journal of Hydrology*, 541, pp.1003-1019.

850 Meredith, E.P., Maraun, D., Semenov, V.A., Park, W., 2015. Evidence for added value of convection-permitting  
851 models for studying changes in extreme precipitation. *Journal of Geophysical Research:*  
852 *Atmospheres*, 120(24), 12500-12513.

853 Monteith, J.L., 1965, July. Evaporation and environment. In *Symp. Soc. Exp. Biol*, 19, No. 205-23, p. 4).

854 Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models. Part I: a discussion of  
855 principles. *J. Hydrol.* 10, 282-290.

856 Nguyen, H., Mehrotra, R., Sharma, A., 2017. Can the variability in precipitation simulations across GCMs be  
857 reduced through sensible bias correction?. *Climate Dynamics*, 1-19.

858 Onöl, B., 2012. Effects of coastal topography on climate: High-resolution simulation with a regional climate  
859 model. *Clim. Res.* 52: 159-174. doi: 10.3354/cr01077.

860 Osborn, T.J. and Hulme, M., 1997. Development of a relationship between station and grid-box rainfall  
861 frequencies for climate model evaluation. *Journal of Climate*, 10(8), pp.1885-1908.

862 Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., Loumagne, C., 2005. Which potential  
863 evapotranspiration input for a lumped rainfall-runoff model? Part 2 – Towards a simple and efficient  
864 potential evapotranspiration model for rainfall-runoff modelling. *J. Hydrol.* 303, 290-306. doi:  
865 10.1016/j.jhydrol.2004.08.026.

866 Parana Manage, N., Lockart, N., Willgoose, G., Kuczera, G., Kiem, A.S., Chowdhury, A.F.M., Zhang, L. and  
867 Twomey, C., 2016. Statistical testing of dynamically downscaled rainfall data for the Upper Hunter  
868 region, New South Wales, Australia. *Journal of Southern Hemisphere Earth System Science*, 66(2), 203-  
869 227.

870 Pasten Zapata, E., 2017. An assessment of the impact of climate change on the efficiency and feasibility of four  
871 run-of-the-river hydropower schemes in UK catchments (Doctoral dissertation, University of Sheffield).

872 Piani, C., Haerter, J.O., Coppola, E., 2010. Statistical bias correction for daily precipitation in regional climate  
873 models over Europe. *Theor. Appl. Climatol.* 99, 187-192. doi: 10.1007/s00704-009-0134-9.

874 Prein, A.F., Gobiet, A., Truhetz, H., Keuler, K., Goergen, K., Teichmann, C., Fox Maule, C., van Meijgaard, E.,  
875 Déqué, M., Nikulin, G., Vautard, R., Colette, A., Kjellström, E., Jacob, D., 2015. Precipitation in the  
876 EURO-CORDEX 0.11° and 0.44° simulations: high resolution, high benefits? *Clim. Dyn.* doi:  
877 10.1007/s00382-015-2589-y.

878 Prudhomme, C. and Davies, H., 2009. Assessing uncertainties in climate change impact analyses on the river  
879 flow regimes in the UK. Part 1: baseline climate. *Climatic Change*, 93(1-2), 177-195. doi: DOI  
880 10.1007/s10584-008-9464-3

881 Prudhomme, C., Haxon, T., Crooks, S., Jackson, C., Barkwith, A., Williamson, J., Kelvin, J., Mackay, J., Wang,  
882 L., Young, A., Watts, G., 2013. Future flows hydrology: an ensemble of daily river flow and monthly  
883 groundwater levels for use for climate change impact assessment across Great Britain. *Earth Syst. Sci.*  
884 *Data* 5: 101-107. doi:10-5194/essd-5-101-2013.

885 Qian, B., Wang, H., He, Y., Liu, J., De Jong, R., 2015. Projecting spring wheat yield changes on the Canadian  
886 Prairies: effects of resolutions of a regional climate model and statistical processing. *Int. J. Climatol.* doi:  
887 10.1002/joc.4571.

888 Robinson, E., Blyth, E., Clark, D., Comyn-Platt, E., Finch, J., Rudd, A., 2017a Climate hydrology and ecology  
889 research support system meteorology dataset for Great Britain (1961-2015) [CHESS-met] v1.2. NERC  
890 Environmental Information Data Centre 7 February 2017, [https://doi.org/10.5285/b745e7b1-626c-4ccc-](https://doi.org/10.5285/b745e7b1-626c-4ccc-ac27-56582e77b900)  
891 [ac27-56582e77b900](https://doi.org/10.5285/b745e7b1-626c-4ccc-ac27-56582e77b900)

892 Robinson, E.L., Blyth, E.M., Clark, D.B., Finch, J., Rudd, A.C., 2017b. Trends in atmospheric evaporative  
893 demand in Great Britain using high-resolution meteorological data. *Hydrology and Earth System*  
894 *Sciences*, 21(2), 1189.

895 Rockel, B., 2015. The regional downscaling approach: a brief history and recent advances. *Curr. Clim. Change*  
896 *Rep.* 1:22-29. doi: 10.1007/s40641-014-0001-3.

897 Rockel, B., Will, A., Hense, A., 2008. Special issue on regional climate modelling with COSMO-CLM (CCLM).  
898 Meteorol. Z. 17, 477-485.

899 Rummukainen, M. B., 2016. Added value in regional climate modeling. WIREs Clim. Change 7:145-159. doi:  
900 10.1002/wcc.378

901 Rummukainen, M.B., Rockel, L., Bärring, J., Christensen, J.H., Reckermann, M., 2015. 21st century challenges  
902 in Regional Climate Modeling. Bull. Amer. Meteor. Soc. doi:10.1175/BAMS-D-14-00214.1.

903 Samuelsson, P., Jones, C.G., Willén, U., Ullerstig, A., Gollvik, S., Hansson, U., Jansson, C., Kjellström, E.,  
904 Nikulin, G., Wyser, K., 2011. The Rossby Centre Regional Climate Model RCA3: model description and  
905 performance. Tellus 63A:4-23. doi: 10.1111/j.1600-0870.2010.00478.x.

906 Seiller, G. and Anctil, F., 2016. How do potential evapotranspiration formulae influence hydrological projections?.  
907 Hydrological Sciences Journal, 61(12), 2249-2266.

908 Scharffenberg, W.A., 2013. Hydrologic Modeling System HEC-HMS User's Manual Version 4.0. Hydrologic  
909 Engineering Center, US Army Corps of Engineers; 442pp. Davis, USA.

910 Skamarock, W.C., Klemp, J., Dudhia, J., Gill, D., Barker, D., Wang, W., Powers, J., 2008. A description of the  
911 advanced research wrf version 3, Tech. Rep., NCAR Technical Note 475, 113pp.

912 Tanguy, M., Dixon, H., Prosdociimi, I., Morris, D. G., Keller, V. D. J. (2014). Gridded estimates of daily and  
913 monthly areal rainfall for the United Kingdom (1890-2012) [CEH-GEAR]. NERC Environmental  
914 Information Data Centre. <https://doi.org/10.5285/5dc179dc-f692-49ba-9326-a6893a503f6e>

915 Teng, J., Potter, N.J., Chiew, F.H.S., Zhang, L., Wang, B., Vaze, J., Evans, J.P., 2015. How does bias correction  
916 of regional climate model precipitation affect modeled runoff? Hydrol. Earth Syst. Sci. 19, 711-728. doi:  
917 10.5194/hess-19-711-2015.

918 Teutschbein, C., Seibert, J., 2012. Bias correction of regional climate model simulations for hydrological climate-  
919 change impact studies: Review and evaluation of different methods. J. Hydrol. 456, 12-29. doi:  
920 10.1016/j.jhydrol.2012.05.052.

921 Tolika, K., Anagnostopoulou, C., Velikou, K. and Vagenas, C., 2016. A comparison of the updated very high  
922 resolution model RegCM3\_10km with the previous version RegCM3\_25km over the complex terrain of  
923 Greece: present and future projections. Theoretical and applied climatology, 126(3-4), pp.715-726.



924 Tölle, M.H., Schefczyk, L. and Gutjahr, O., 2017. Scale dependency of regional climate modeling of current and  
925 future climate extremes in Germany. *Theoretical and Applied Climatology*, pp.1-20.

926 Torma, C., Giorgi, F., Coppola, E., 2015. Added value of regional climate modeling over areas characterized by  
927 complex terrain – Precipitation over the Alps. *J. Geophys. Res. Atmos.* 120:3957-3972. doi:  
928 10.1002/2014JD022781.

929 van Meijgaard, E., van Ulft, L.H., Lenderink, G., de Roode, S.R., Wipfler, L., Boers, R., Timmermans, R.M.A.,  
930 2012. Refinement and application of a regional atmospheric model for climate scenario calculations of  
931 Western Europe. *Climate changes Spatial Planning publication: KvR 054/12*.

932 Velázquez, J.A., Troin, M., Caya, D., Brissette, F., 2015. Evaluating the time-invariance hypothesis of climate  
933 model bias correction: implications for hydrological impact studies. *J. Hydrometeor.* 16, 2013-2026. doi:  
934 10.1175/JHM-D-14-0159-1.

935 Walsh, C. L., Blenkinsop, S., Fowler, H. J., Burton, A., Dawson, R. J., Glenis, V., Manninga, L.J. and Kilsby, C.  
936 G., 2015. Adaptation of water resource systems to an uncertain future. *Hydrology & Earth System  
937 Sciences Discussions*, 12(9).

938 Wang, H. M., Chen, J., Xu, C. Y., Chen, H., Guo, S., Xie, P., & Li, X., 2019. Does the weighting of climate  
939 simulations result in a better quantification of hydrological impacts?. *Hydrology and Earth System  
940 Sciences*, 23(10), 4033-4050.

941 Wetterhall, F., Pappenberger, F., He, Y., Freer, J. and Cloke, H.L., 2012. Conditioning model output statistics of  
942 regional climate model precipitation on circulation patterns. *Nonlinear Processes in Geophysics*, 19(6),  
943 pp.623-633.

944 Wilby, R. L., and Harris, I., 2006. A framework for assessing uncertainties in climate change impacts: Low-flow  
945 scenarios for the River Thames, UK. *Water Resources Research*, 42(2).

946 Wilcke, R.A.I., Mendlik, T., Gobiet, A., 2013. Multi-variable error correction of regional climate models. *Climatic  
947 Change*. 120, 871-887. doi: 10.1007/s10584-013-0845-x.

948 Yang, W., Andréasson, J., Graham, L.P., Olsson, J., Rosberg, J., Wetterhall, F., 2010. Distribution-based scaling  
949 to improve usability of regional climate model projections for hydrological climate change impacts  
950 studies. *Hydrology Research*, 41(3-4), 211-229.

951 Zittis, G., Bruggeman, A., Camera, C., Hadjinicolaou, P., Lelieveld, J., 2017. The added value of convection  
952 permitting simulations of extreme precipitation events over the eastern Mediterranean. *Atmospheric*  
953 *Research*, 191, 20-33.

954